1. Specification of a conversational domain

A conversational domain, R, can be constructed with respect to an arbitrary subject matter; probability theory, history, chemistry or geography. It may be constructed in several ways, one of which is described in the present chapter; namely, by a controlled natural language dialogue involving teachback; which takes place between a human source of subject matter and a human "interrogator analyst".

The "interrogator analyst" translates loose natural language expressions about topics (designated T_i) into precise ones about relations R_i, expressed in L^*. The R_i can be manipulated using relational operators. Later, when the manipulated structure has been tested for various conditions (consistency, cyclicity, and so on) portions of it are discussed with the source; taught back, matched and modified, usually in terms of a loose retranslation using the source idiom.

The description of source/interrogator analyst dialogue represents a compromise between the demands of pragmatism and rigour. Both qualities suffer to some extent, though not irremediably. From a practical point of view, the chief inaccuracy is due to couching the process in orderly terms; in reality, many events, presented as quasi-serial, take place in reverse sequence or even simultaneously. So far as rigour is concerned, the translation process has no algorithmic pretension and consequently not everything can be precisely specified. When exact definitions are possible (e.g. of relational operators, criteria for closure and the like) they are interpolated at appropriate points (which keeps some unfamiliar thoughts in context and avoids an indigestible bolus of uninterpreted symbolism at the outset).

1.1. *Dialogue.* Let L^+ be an unrestricted natural language. It need differ from L^* only insofar as L^+ expressions are not subject to the kind of logical scrutiny and justification used, for example, by Reichenbach (1947), whereas dialogue in L^* definitely is liable to such an analysis. The L^+ dialogue takes place between a source and an "interrogator analyst". Their dialogue is a conversation but clearly it is not a strict conversation and consequently "conversation" will be suppressed (to avoid prefixing nearly every other occurrence of this term by "strict"). It is essential that the source as well as the interrogator analyst should be able to understand L* though, as a practical expedient, the source is rarely if ever called upon to do so. Most of the L* statements in the dialogue appear as L^+ translations.

The interchange between the source and the IA (interrogator analyst) has a domain itself; namely, a body of subject matter the source knows about. Moreover, there is an intention on the part of both the source and the IA to construct the (formal) domain for a strict and usually tutorial conversation between other people.

1.1.1. Various sources have been used for this purpose; for example, subject matter experts, skilled craftsmen, faculty members, curriculum designers, a design team, knowledgeable critics with access to books and other literature about the subject matter.

1.1.2. Various subject matters have been treated in this manner; for example, logic; history, regarded as a role playing in period; history, regarded as the art of finding evidence to support an hypothesis; segments of physical science; statistics and probability theory viewed as an applied science; parts of biology and bio-chemistry involving cyclic processes such as the operon-repressor system and the mammalian oestrous cycle, map-reading and various laboratory sized taxonomies.

1.1.3. The IA (interrogator analyst) is someone who knows the purpose of the dialogue and is also versed in the analytic rules and questioning techniques indicated in the sequel. It is helpful if the IA has computing facilities and certain graph drawing devices at hand and he may need various psychological tools such as forms for administering repertory grids.

1.2. *Technique.* Like an interviewer doing motivational research or a psychotherapist interacting with a patient, the IA has to face different stances and orientations of the source; some of them perverse. He can reasonably assume that the source desires to reveal his knowledge and will chiefly speak about the given subject matter. The entire dialogue is prefaced by a statement like "tell me all the conceivable ways you could teach what you know of

the subject matter" provided only each topic is somehow connected (say how) to some other topic. (The say how part is particularly important. At some stage the IA must be able to formulate connections, that is, relations, between every topic cited by the source.)

1.2.1. With luck, this evokes a stream of free association about the subject matter and how it might be represented, i.e. the citation of topics and their connections. From time to time, however, the stream is narrowed (the source often has an almost unconquerable inclination to state one way of teaching something, rather than "all the ways he can conceive") and, if so, an IA must ask "are there any other ways" and possibly probe the source's intellect. Occasionally the flow of dialogue dries up completely. If so, the IA assumes that the source is basically willing to reveal what he knows but is, for one of various reasons, unable to do so in practice. For example, there are virtually limitless reasons for preserving esoteric wisdom. They would be sound enough reasons, for instance, in the context of a consultation with a client. Few of them have any bearing on this situation. Even so, the mental mechanisms for defending private concepts (used legitimately in the professional field) come into play automatically and it is up to the IA to thwart these manoeuvres by exposition or questioning. Again it may be that the source is simply unable to express (in L^+) what he actually comprehends. If so, the IA prods the source with enquiries or, if the blockage is resilient, he can resort to Kelly's (1955) repertory grid technique, using previously stated or well known topics as "objects" and eliciting source concepts as "constructs" (the fresh topics generated by the source concepts being possible attributes of all other topics not just the "object" topics used to elicit them). It is often appropriate to employ Hinkle's "laddering" method or one of the comparable devices considered by Bannister and Mair (1968) to extract "superordinate" and "subordinate" constructs, as well (so far as an IA is concerned, just "related topics").

1.2.2. No explicit bounds are placed upon the scope of knowledge. For example, the source may exhibit chemistry in text book terms, in terms of its historical development or as the choreography of a ballet of molecules. He can talk of chemistry in any way he likes or in any way he believes a student might think of it; the more the better. Expanding the scope of discourse is rarely troublesome except in connection with procedural and perceptual motor skills where expertise is typically maintained by citing arcane principles or abilities to be acquired only by experience (in apprenticeship or the like).

1.2.3. Nor is there any truth requirement. For example, in the subject matter field of chemistry, the source might describe the phlogiston theory (which is false, but perfectly teachable) or even alchemy (which hangs together in its own fashion).

2. Constraints on dialogue and teachback

Certain ground rules are imposed upon this free wheeling discourse to secure an L^* translation of the source's knowledge. The rules are applied continually as condition tests upon a teachback situation of the type suggested by Fig. 1. When the source has cited a few topics (T_i) the IA builds up his tentative L^* translations (R_i) and represents his translation of the relations asserted to hold between the T_i in terms of a net, Ω of which the R_i are part. This net consists of a list $\{ R_i \}$ of the R_i currently held to be in (tentative) correspondence with the T_i and of relational operators (relations between relations) described in L^*.

From time to time the IA's L^+ translation of the currently existing Ω is taught back to the source who may or may not approve of it as a veridical statement of his knowledge. If not, Ω is modified to bring the source's image and the IA's image into register.

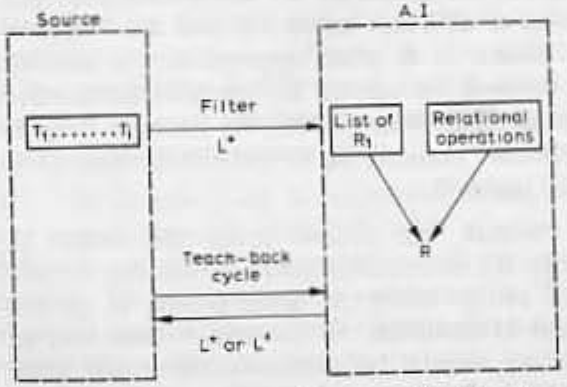


Fig. 7.1. Overview of translation, teachback, and matching.

2.1. *Rules.* There are five rules. *Rule 1* is that any topic T_i cited in L^+ has an image R_i in L^*. This rule is a postulate about knowledge, engendered by the present theory, and henceforward assumed. *Rule 2* is that any cognitive transformation, reported in L^+, can be translated into an L^* transformation carrying relations R_j, R_i into a relation R_i and made up from a standard collection of relational operators that are used as part of an L^* derivation. Though plausible in its own right, Rule 2 entails the strong theoretical postulate that there are concepts <u>Proc</u>^{1} in any cognitive repertoire, π, for realising all of the relational operators. Previously, this postulate was referred to a student and introduced in order to predict that any student can learn certain L^0 concepts <u>Proc</u>^{0}. Here, it is referred to the source and is formulated as a translatability condition. The IA uses relational operators to manipulate R_i and to build Ω from its component R_i. By hypothesis (hereafter assumed) such operators are intelligible to the source also. Thus the IA can furnish an L^+ translation of what a relational operator is or what it does when applied to a cluster of relations. If necessary, he can teach the definitions given in Section 3 and 4 to the source. *Rule 3* is that a topic T_i is accepted if and only if it is related to some other topic T_j. That can either be regarded as a postulate about knowledge ("knowables are connected") or an acceptance criterion. *Rule 4* calls for consistency; different topic relations have different names and the same relation the same name. This rule is an acceptance criterion for topic names, required in order to avoid unintended ambiguity and loss of specificity. It also underpins rule 5, insofar as rule 5 is applicable to relations accepted under rule 4. Finally, *Rule 5* is that a connected list of topic relations is a valid candidate for a conversational domain only if the relations between the relations are cyclic in form. A condition test (for cyclicity) is applied to the IA's L^* construct of Ω, or part of it, whenever the source asserts that R (the domain) is part or all of Ω.

2.2. *Implementation.* The following section gives an account of how these rules are applied in the course of dialogue between the source and the IA. Though the IA is responsible for ensuring that the rules are obeyed (he acts as an umpire or an arbitrator) it is assumed that the source is aware of the rules and generally respects them.

2.2.1. *Rule 1.* "A source has knowledge" means, "the source has reproducible concepts which (as previously) are procedures for computing/instantiating/bringing about whatever topics T_i are cited". T_i in other words, is a relation expressed intensionally in L^+ and the index "i" of "T_i" is the name for a source concept that brings about T_i.

In particular, the source can compute his extension of T_i (i.e. the listing of whatever T_i relates) and this property is crucial to the argument because the IA's L^* image of T_i, namely R_i, is manipulated (by relational operators) as though in extension (in a partially or completely listed form).

2.2.1.1. Informally, designate the source operation which exhibits his extension of T_i "unzipping". It goes as follows. T_i is cited, Unzip T_i (to yield the entities it relates, say T_x, T_y). Now T_x, T_y could be noun-like objects, but, in general, they are other relations (consider the unzip of Perception, for example). If so, there are also, by hypothesis, source concepts x, y, to compute the relations T_x, T_y and the process can be repeated to give (say) unzip $T_x = T_a, T_b, T_c$ and unzip $T_y = T\alpha, T\beta$. No limits are imposed upon this process excepting those immanent in the source i.e. the IA does not demand that the related entities are unitary, in the sense that they are really elements of unordered sets. Whatever the IA and the source agree to count as unitary during teachback, will serve well enough (though if Ω is to satisy *Rule 5* and if its constituent R_i are to satisfy *Rule 3* some constraint is imposed upon the grain of these entities).

2.2.1.2. The formal operation which images unzip in L^* is called Field; that is, for R_i (a translation of T_i)

Field $(R_i) = <\text{Domain } R_i, \text{Range } R_i>$, (since R_i is a relation).

But the domain and the range of R_i may both be relations R_x, R_y (if not, they are called simple) so that

$$\text{Field (Field } (R_i)) = << \text{Domain } (R_x), \text{Range } (R_x)> \\ <\text{Domain } (R_y), \text{Range } (R_y)>>$$

Repeating the process (without limit unless a field is simple) yields Field Iterate (R_i) which is an ordered and indexed set of relations (for example, the terms of Field (Field (R_i)) may be written using notations like Field (Field (R_i)) $= R^i_{11}, R^i_{12}; R^i_{211}, R^i_{212}, R^i_{221}, R^i_{222}$. Taken alone, this exercise simply generates

vacuous index sets[16]; the IA does not know, by virtue of applying Field Iterate to separate relations R_i, R_j, whether or not there are interesting correspondences such as $R^j_{km} = R^i_{in}$. But the process does generate place holders; that is, blank and labelled spaces in an L^* description, to be filled under translation with the entries produced (by the source, in L^+) if he *unzips* T_i.

2.2.1.3. Using the broad arrow to stand for translation and construing the IA as calling for an unzip after completing any one stage in the Field operation the $\overline{\text{(least)}}$ non trivial interaction is shown in Fig. 2.

2.2.2. *Rule 2.* All L^* relational operators (a set of them is specified in Section 3) can be represented in L^+.

The relational operators (discussed more precisely in Section 3 and 4) are abstract relations between relations. For example, projection (as in the projection of points in a plane onto one coordinate, or projection of points in a higher dimensional space onto a plane) is a relational operator that acts upon one relation R_i: composition (as in the composition of functions or operations) is an operator that acts upon a cluster of relations.

According to *Rule 2* any operator, O_X, or any sequence of operators, O_Y, can be expressed, in context, as an L^+ translation. The circumstances under which an L^+ translation of L^* operators is important are as follows: The source (either of his own volition or because he is required to do so), asserts that a topic T_j is

[16] The relational operator Field Iterate is not so simple minded as it seems to be. Let R be a relation. Taken in extension R_i is a subset of some (undefined as yet) cartesian product set of the necessary coordinates of R_i. First (since the number of coordinates is so far unstated) we generate numbers of them and numbers are ordered in some manner. This process requires a Distinction or Predication; a base (say "1"); and a successor operation to generate integers and a permuting operation to order any finite number of coordinates. Next, R_i is projected onto coordinates which are named (if necessary after permutation) by the source (as "X" or "Y" in the example cited). If the elements of the projection Proj[j] (R_i) onto the jth coordinate are simple (in the set theoretic sense) nothing more is needed to specify this fact. In general, however, the elements of Proj[j] (R_i) are relations (a, b, c, or α, β in the example); in fact, counting a property as a 1 adic relation, I believe they always are. So, a further distinction is needed and a further index (of elements of the "1" fields) must be generated under successor. Of the requisite operations successor and permute and project etc. are later defined as relational operators; Distinguish (or predicate) is formally simulated.

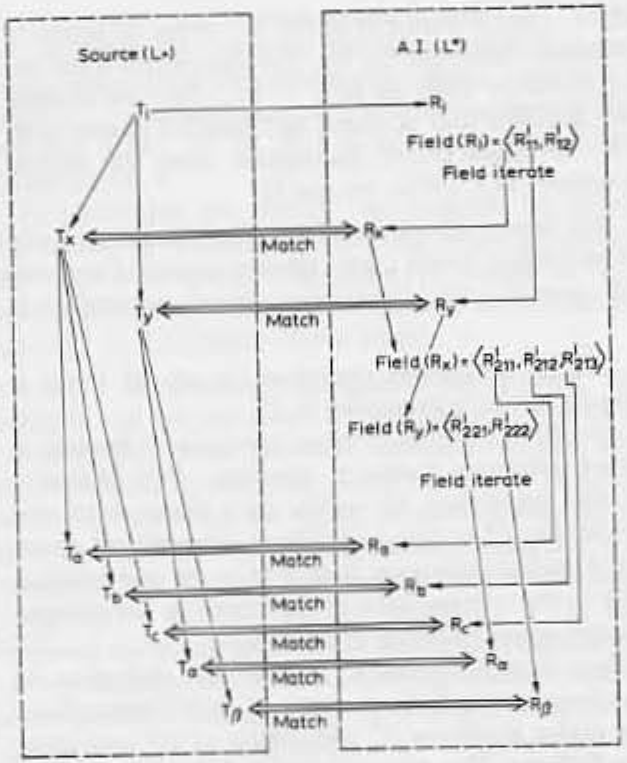


Fig. 7.2. Unzipping by teachback, where unzip is seen to consist in field iterate and match whilst maintaining consistency of names.

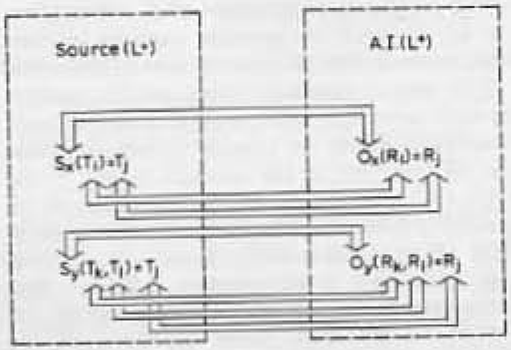


Fig. 7.3. Connectivity translation.

somehow derived from one topic π_i, or several T_i, T_j. These topics are imaged in the IA's L* image as abstract relations R_i, R_j and R_i. The IA proposes that $R_j = O_X(R_i)$ or that $R_j = O_Y(R_i, R_i)$ and is able to give a context specific L⁺ translation of this contention; for example, ($T_j = S_X(T_i)$) "Distance travelled is the time integral of velocity" or ($T_j = S_Y(T_i, T_i)$), "Social homeostasis is an equality relation between role occupancy and population density". Since S_X and S_Y are context specific there are generally several S_X (say $S_X{}^1$, $S_X{}^2$, ...) to each O_X: several S_Y (say $S_Y{}^1$, $S_Y{}^2$, ...) to each O_Y. Of these, the source may select any one (and only one) as being the case, i.e. as being a valid translation of the association between T_i and T_j that he has in mind; if so, then (Fig. 3), S_X matches O_X or S_Y matches O_Y. If none of the S_X or S_Y are the case then there is no match. Conversely, the source may cite S_X or S_Y, and insofar as this derivation corresponds to an abstract operator or sequence of operators O_X or O_Y, the IA will accept it unless it leads to a contradiction of previous statements.

The contention is that any O_X or O_Y can be given an L⁺ translation S_X or S_Y. As a conjecture, any possible derivation or explanation in L⁺ also corresponds to an L* derivation O_X or O_Y.

2.2.3. *Rule 3.* Topic relations are connected. For each topic cited by the source, he should be able to establish some relation between the topics (trivially, this is always true; he cited the topics in order!). In practice, the IA is interested in the relations that the source wishes to emphasise.

As a consequence of this rule the R_i in $\{R_i\}$ (the IAs L* list) are connected to form a net, Ω. If T_i and T_j are cited by the source there is at least one relation T_o between them, i.e. $T_i\ (T_o)\ T_j$. Let all topics be translated, suppose T_j is cited after T_i. At the moment T_j is cited there are two distinct relations R_i and R_j. Rule 3 insists that the two representations are reduced to one; namely $T_i\ (T_o)\ T_j$ which means (as in Fig. 3) that $R_j = O_X(R_i)$ or $R_j = O_Y(R_i, R_i)$.

It does not follow that there is just one relation between R_j and R_i from which the cited relation (R_j) is uniquely derived. In general, there are many relations, and it is usually necessary to countenance several of them. For example, (in a medical course) the relation "diagnosis of cells in the thyroid gland producing excess of thyroxine and or triiodotyrosine" is derivable from a relation between the cell count and cell size of the thyroid gland, called "Toxic diffuse goitre", or it is derivable as a behavioural relation between certain regulatory systems characterised by overgrowth and overactivity of the cells in this gland.

To ensure that connectedness exists, the topics cited are held in abeyance, momentarily, by the filter of Fig. 1 and T_j for example, is accepted if and only if T_o is elicited.

2.2.4. *Rule 4.* Recall that the name of topic T_i is i. Consistency (the mandate of this rule) requires the source to name topics in such a way that one and only one name is given to each topic; the same topic is never given a different name and no name is assigned to more than one topic. Authors frequently ignore or disobey this rule (for example, over a sample of more than 60 standard statistics text books the name "distribution" was used to denote, on average, 8 different topics).

The test for consistency is applied whenever any topic T_j is placed in correspondence with some topic relation R_j. The test starts by a simple match of names, j, between Ts and Rs. If it detects a mention of the same relation (i.e. some relation R_j named "j" exists in the list $\{R_i\}$), then consistency requires that the name, j, refers to the same relation. If T is a fresh topic (that is, no relation R_j named "j" existed in the list) consistency requires that R_j is not the same as any other relation. Due to our construction the equality "some relation" is ambiguous. "Some" might mean that R_j is isomorphic to some other relation $R_j^{\square}$ (even though R_j and $R_j^{\square}$ have different fields in one to one correspondence) or it might mean that $R_j = R_j^{\square}$ (in the sense that R_j is a subset of m-tuples and $R_j^{\square}$ is the same subset of the same m-tuples). This ambiguity is resolved by the following considerations. If the network Ω is (at a later stage) accepted as a valid conversational domain, R, then it will be given at least one L^1 description $D^1(R)$ in which each topic relation R_i in R is uniquely specified. Moreover, each R_i in R will be given an L^0 description $D^0(R_i)$ (canonically a command graph which stipulates how, according to the source's theory, R_i may be brought about). Either $D^0(R_i)$ consists in a class of L^0 explanations, given under an interpretation of L^0 that is chosen by the source or else (the case we are primarily concerned with) $D^0(R_i)$ is a class of models, constructible in a given modelling facility, which explains R_i in the context of this facility, insofar as R_i is brought about if any one of them is executed.

In either case, we require that if $R_j = R_j^{\square}$ then $D^0(R_j) = D^0(R_j^{\square})$ and that R_j, having a unique L^1 description $D^1(R_j)$ corresponds to one and only one $D^0(R_j)$.

If an inconsistency is detected, it may be due to a one to many or a many to one correspondence between names, j and $D^0(R_j)$.

If there is a many to one defect then T_i refers to R_i in one context and T_j to R_i, in another. This usage is disallowed unless the source can show that R_i is equivocal (i.e. that it should be replaced by a pair R_j and R_k) in which case the requisite adjustment is made. Otherwise the many to one defect is remedied by deleting one of T_j, T_k, or by deleting both T_j and T_k when the source must invent a new topic name.

A one to many defect is manifest as the assignment of one topic name to two or more distinct topic relations; for example "motor car" (a topic name) is used to designate "motor car as understood by a mechanic" (in L*, a mapping between certain input/output properties of the engine, electrical system, and so on) and "motor car as understood by the driver" (in L*, a mapping between road configurations and control parameters). Obviously, the two senses of "motor car" are related (for instance, when examining skid avoidance, they have a non empty intersection and in the universe of vehicle control characteristics they are isomorphic). But the two relations are distinct because "motor car" was declared in different contexts; it was not declared either as the intersection manifest in "skid avoidance" or in the universe of "vehicle control" as an isomorph. This kind of defect is remedied by introducing fresh topic names that are in register with the existing R_i; usually, the temporary names from an arbitrary alphabet. For example, "Motor Car" becomes either MC_A or MC_B to be replaced later by something more elegant. Meanwhile, MC_A and MC_B are used as labels on the corresponding L* coded representation which is periodically taught back to the source.

2.2.5. *Rule 5* asserts that any conversational domain R (with topic relations R_i in R) is a network in which the R_i are related in a cyclic fashion[17]. From time to time the source is encouraged to nominate some head topic, T* in the nexus of topics he has in mind (which picks out some relation R* in the IA's network Ω) and to submit Ω, or part of it, as a candidate domain. There are no restrictions imposed upon when the source makes his submission, but, whenever he does so, the IA displays (most conveniently, as a graphic representation) an image derived from the current network Ω in which all of the R_i (in list $\{ R_i \}$) are distinct and in which the relation (R*), corresponding to the head topic (T*), is singled out for attention. The source is allowed to circumscribe the relations around R* that he wishes to consider as part of a conversational domain, and the ritual employed for this purpose involves a notion (a vector, Δ, of "depths") which is explained in Sections 5 to 8. At this point, the possibly circumscribed Ω is tested for cyclicity using a fully mechanisable procedure, also described in Sections 5 to 8, but designed to secure the following condition, for R* and any R_i, R_j in Ω that are used in deriving R.

$$R^* = O_u(R_i, R_j): R_i = O_v(R^*, R_j): R_j = O_Y(R^*, R_i)$$

That is, if R* is derived, though O_u, from R_i and R_j then R_i may be derived by O_v, from R* and R_j whilst R_j can be derived, by O_Y, from R* and R_i. If the source is a teacher, this requirement seems patently reasonable. R* is explained in terms of R_i and R_j,

[17] A much stronger notion of cyclicity is developed (chiefly due to Kallikourdis) in the next volume. Entailment meshes are reduced to conjunctive substructures related by morphisms. It is shown that a conjunctive substructure is strongly cyclic in the sense that it can be reconstructed, or reproduced, without recourse to storage of any but the topic relation at its head if its relations are related either by the relational operator natural join only or by natural join and union (provided there is a chain of natural joins from each relation in the structure to the head). The reproducing program for any such structure involves the relational operator's partial complementation, union and projection. An arbitrary conjunctive substructure is strongly cyclic if it has the form outlined, or if it is possible to apply a transformation that replaces any projection in the original by identity and any restriction or composition by natural join (leaving union unchanged). Unless the arbitrary structure contains unpropitious unions the transformed structure is strongly cyclic and an inverse transformation that augments the reproductive program, carries it back into the original.

which are both subordinate to it, if the subject matter is viewed with R* at the head; any subordinate topic relation of R* should be derivable from R* using no more topic relations than the other subordinates needed in order to reach R*. For example, in molecular biology, if the topic relation "operon" is cited as head and is derived from the subordinates "repressor" and "operator gene" amongst others, then "repressor" should be derivable from "operon" if "operator gene" is given (together with the other topic relations needed to establish "operon"), and "operator gene" should be derivable from "operon" if "repressor" is given. In psychological parlance a cyclic Ω is a Gestalt; epistomologically, it corresponds to a coherent theory entertained by the source (not necessarily a valid theory; for example modern chemistry and phlogiston theory are both acceptable on these grounds).

2.3. If Ω, submitted under a head R*, is found to be cyclic or to contain a cyclic component then it, or its cyclic component, is accepted as a conversational domain, R; failing which, Ω's deficiencies are pointed out to the source. Suppose Ω is accepted as R, certain simplifying operations called pruning are carried out and the pruned form of R is displayed.

It should be emphasised that if Ω has a cyclic component under any one head then it usually has cyclic components under several heads. For example, the relations in our main example (probability theory) are cyclically connected under this head and under "probabilistic automata" and many others. The choice of head does not influence the underlying Ω, but the pruned form depends crucially upon the head. In fact, choice of a head is the first step in giving an L^1 description, $D^1(R)$, of the putative knowables.

After R is accepted (and its pruned form is displayed) the source is at liberty to extend the domain by constructing further topic relations and to express their relation to one another and to the topic relations that exist in Ω. In turn, a further head may be chosen and the cyclicity condition tested as an iteration of the process already discussed. It does not follow, of course, that the new Ω will be cyclic and it may be true that the old Ω has become non-cyclic as a result of adding further topic relations.

2.4. Whenever he receives a pruned form of R (as he does whenever Ω and a head turn out to be cyclic) the source may furnish an L^1 description, $D^1(R)$, of R and an L^0 description $D^0(R_i)$ of how to bring about each R_i in R. The essential constraint is that he must do so before finalising R and most sources find it convenient to keep their descriptions updated.

The detailed requirements for $D^1(R)$ are described in Section 4, and so is the form of $D^0(R_i)$. By way of orientation $D^1(R)$ is a (possibly redundant) indexing scheme, in which the topic relations are unambiguously located; $D^0(R_i)$ may either be a set of L^0 explanations of how to bring about R_i in a field of interpretation chosen by the source, or it may be a class of models built, in a given modelling facility, that demonstrate R_i insofar as R_i is brought about if any of these models are executed.

2.5. The iterated process satisfying *Rule 1* to *Rule 5* is represented by the teachback cycle of Fig. 1. From Rule 1, any T_i corresponds to some R_i which may be decomposed, and its factors (obtained by "unzip") may also be matched under translation, unless it happens that this topic contravenes Rule 4. From Rule 3 each R_i, R_j cited by a source is related to some T_i, T_j and, from Rule 2, the relation between R_i, R_j can be expressed in terms of L^*, as a relational operator O_X or a sequence O_Y; such L^* structures are matched, under translation, to S_X or S_Y in L^+. Either kind of matching, of factors or constructions (and, if the latter, constructions due either to the source or the IA) constitutes teachback i.e. matching explanations with re-explanations. During the teachback operation Ω is tested for cyclicity under Rule 5 to yield (eventually) R as a domain. R is described in L^1 (as $D^1(R)$) and all R_i in R are described, in L^0, as $D^0(R_i)$.

2.6. The process is shown as a flow chart in Fig. 22. This chart cannot be fully appreciated until some further terms are introduced. It is, however, intelligible, even at this stage and the reader may find it useful to glance at the chart, in order to obtain an overview.

The following slightly repetitious comments lend meaning to the flow charted procedure. It will be evident that a source's topics (T_i) are semantically interpreted relations; albeit Fuzzy or underspecified relations. They are interpreted in any universe or class of universes that accommodates a verbal or non-verbal explanation (for example, the real universe of mechanical processes and the objects upon which they act; the intellectual universe of mathematical entities). In contrast the AI manipulates

relations in extenso, devoid of specific interpretation; these are the R_i. At the stage of matching (T_i to R_i) he needs only consider an interpretation scheme sufficient to establish consistency and thus he leaves open many universes of interpretation which might be legitimately espoused by a student; all the same, he does instate certain distinctions between classes of universe of interpretation, namely, those distinctions essential to his thesis.

Later on, a full semantic interpretation is retrieved and adjoined to the network of (extensional) relations R_i, by two operations (the two lowest boxes in Fig. 22). First, the source is required to exemplify each topic by citing demonstrations/legitimate verbal/non-verbal explanations of each R_i. These, together with a modelling facility, are described by an L^0 description $D^0(R_i)$ and form the basis for a task structure TS(i). The operation is detailed in Section 9. Next, the source is required to describe the syntactic structure (nodes in a network, acting as place holders for relations) by assigning and evaluating L^1 predicates over the nodes. This operation is detailed in Section 10. The L^1 description, $D^1(R)$, gives a partial semantic interpretation to the syntactic network. In particular, it demarcates classes of universe in which models can be built as non verbal explanations. But the student is still usually allowed a great deal of latitude over the interpretation he does give; the model he builds, the entities upon which it acts when it is executed. In this book, the essential degree of latitude is noted and stressed. In the next volume this feature of a conversational domain is stated in terms of Fuzzy Set Theory.

2.7. There is no well specified end to this process. It might terminate after the source submits the first Ω that passes the cyclicity test and is counted as a valid conversational domain. But there is no reason why the source should stop at this point. In general, the longer construction goes on the more that is explained and/or the more varieties of explanations are assigned to a given topic relation. Provided there is an output of consistent and cyclically related knowables the justification for stopping at a given point is simply that any student in the population intended to learn in this conversational domain would understand the explanations being offered.

There is no logical or epistemological reason for allowing any topics that are not explained i.e. if the source says that something is given (as a definition, for example) he can always be taxed to support it (either by further definitions or by reasons for choosing this particular definition). In practice however, the source is allowed to say that some topics are primitives realising that "primitive" is used at the moment in a relative sense and that later he may be called upon to explain the alleged primitive, after all. Nor is there any logical or epistemological canon for truncating the string of associated topics and names that emanates from the source. But, in practice, the IA does allow the source to stop at whatever the source believes to encompass the subject matter; first asking if that is the only way of adumbrating it (if not, give another) and next cautioning him that it may be necessary if cyclicity is to be secured, to widen the scope, later.

2.8. From the last section, it follows that any source is destined to mark certain relations as primitive unless he continues to expose his knowledge of the subject matter indefinitely. These "primitives" are also bound to be the last fields in the "unzipping" operation (the lowermost elements in Fig. 2). When the source does call a relation primitive he is stating that any student who learns in the conversational domain he has delineated must be able to regard these relations as simple properties i.e. that he has a concept, $\underline{\text{Proc}}^0$, for bringing them about, in his repertoire, π^0. He does not, of course, assert that any student must regard the primitive relations as properties (if he wanted to, the student could dissect the $\underline{\text{Proc}}^0$ into appropriate parts). Even less, is the source asserting that no student regards other relations in the conversational domain as "simple properties"; this will depend upon the student's previous knowledge (to cite the main example, a student already versed in "probability theory" may regard the "least primitive" relation, "probabilistic inference" as a property).

In other words, so far as this theory is concerned, knowledge is endless. The really important feature of a conversational domain is not where the source chooses to stop his exposition, but that, wherever he does stop, he has laid down some Gestalt; some reproducible and knowable cognitive entity. This feature of a conversational domain, R, is guaranteed by the requirement that any R is cyclic. If so, it is knowable and memorable, though the acts of knowing it or remembering it may or may not be within the compass of a particular student. If it happens, however, that the student's repertoire contains $\underline{\text{Proc}}^0$ for all of the primitives and $\underline{\text{Proc}}^1$ that are isomorphic to the relational operators then knowability is certainly also guaranteed. Further, if Rule 2 is accepted, then the latter condition ($\underline{\text{Proc}}^1$ for each relational

operator) is taken to be satisfied for any cognitive repertoire; hence, for any student. This is no more than an assumption, of course, though it appears to be a very reasonable assumption to make of any sentient being.

3. Relations as sets (relations in extensional form) and relations as concepts (in intensional form)

Relations are manipulated by the IA in extension: that is as subsets of the product of their co-ordinates (i.e. the domain, range, pairs in their field) which themselves are sets (disjoint subsets of the co-ordinate sets being values of these factors).

3.1. *Normality*. Standard treatments start off with the assumption that an unordered set is given and structure is imposed upon it (for example, by picking out co-ordinates which are indexed in a given order to form a product). This orthodox picture is inverted by the present approach (as promised in Chapter 5, Section 1) where a concept or procedure or intension is given-as-known and relations are instantiated (by *unzip* and *matching*) as members of the product of the concept's field co-ordinates, taken in an appropriate order. All the same, once that the relation is specified extensionally, it can be manipulated by the usual set theoretic operations.

The source/AI dialogue of the last section is chiefly concerned with getting relations (expressed intensionally as concepts or as procedures) into the extensional form. One representation (a table, with columns standing for the co-ordinates of a Cartesian Product in which the relation is a subset and row entries that indicate values of that co-ordinate) is especially useful and is called (after Codd (1970)) the *Normal* form.

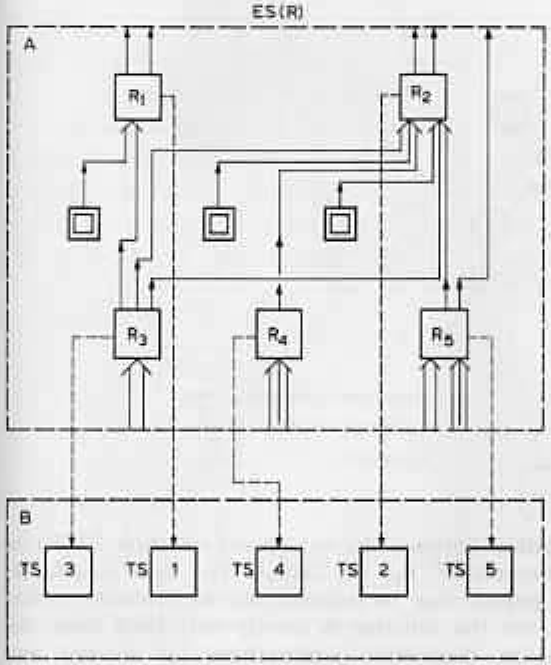
It will be evident that any finite relational network can be represented in normal form using the co-ordinates of its primitives; that is, the relations that any user of this conversational domain must be able to regard as properties because he has concepts in his repertoire that are able to compute these relations. This makes no comment upon the intension of the relations; or the concepts ($\underline{\text{Proc}}_A$ i in the repertoire of a specific student A. Any student (A) may regard R_i as being a property (1-adic-relation), regardless of whether R_i is primitive or not whenever he has a concept $\underline{\text{Proc}}_A^0$ i for computing R_i. Conversely, even if A does have a concept for R_i, he may also analyse the nature of $\underline{\text{Proc}}_A^0$ i (as made up from more elementary subroutines) and, in this case, he does not treat R_i as a property, even though he may do so. The matter is immaterial to manipulative systems, since there is no commitment regarding whether R_i has co-ordinates (1-adic-relations, set of value-sets, for objects ostended by an object-variable) that are really properties or whether they are simply *treated as though* they were.

3.2. *Other examples.* Relational networks are employed quite commonly to represent data and processes for computer programs able to carry out induction, analogical reasoning and the like. Several examples are described in Minsky (Ed.) (1968) and there are many more recent systems of this kind (for instance, Quillian 1969, Winograd 1972). A particularly lucid and non-technical account is given by Winston (1970) in his discussion of pattern recognition with respect to a simple but 3 dimensional environment of building blocks.

A robot receives a (frequently ambiguous) image of the 3 dimensional environment as a 2 dimensional projection onto its "retina". The robots view of the environment is built up in terms of what it can do; either physically or symbolically. In either case, it is ultimately built up in terms of relations between tests (for values of relations) that the robot is designed to count as properties and operations (which are themselves relations). For example, features, such as junctions of detectable lines, are related (by "adjacency" or "right off" and the like) to form figures (cubes and arches and so on); equally, however figures are related to form scenes.

It is useful to think of the relational networks under immediate discussion in much the same way. Their abstract representations are graphs in which nodes stand for relational operators. Somehow, the abstract structure is inscribed and, whenever it is used as the conversational domain of a strict conversation, the inscription is embodied in data storage positions at an interface. The CASTE interface is typical.

3.3. *Inscription at an interface.* In the CASTE interface each storage position has two coupled locations; the coupling is shown in Fig. 4(a) by a dotted line. Each node in the relational network bears a name i, which is inscribed in this location together with an account of how the topic relation R_i is derived from different topic relations R_j in R and a register to indicate the state of the



A = a fragment of $D^1(R)$ (its concrete form is ES(R))
B = a fragment of $D^0(R)$ (its concrete form is TS(R) and is made up from $TS(R_1), TS(R_2), TS(R_3), TS(R_4), TS(R_5)$)

Fig. 7.4 (a). Association between $D^1(R)$ and $D^0(R)$ or, equivalently, ES(R) and TS(R) (repetition of Figure 4.5).

node i of R_i. $D^1(R)$ is an L^1 description of R in which each of the R_i is uniquely discriminated; together with an account (in terms of the relational operators) of how R_i is derived. During any conversation the state of this node (understood, or goal or aim) is delineated by L^1 marker predicate values. The physical inscription of $D^1(R)$ is ES(R).

The remaining member of each coupled storage position in the interface is a link addressing one and only one L^0 description of how R_i may be modelled, explained, and brought about, namely $D^0(R_i)$. The physical embodiment of $D^0(R_i)$ (canonically, a command graph) is a Task Structure, $TS(R_i)$.

3.4. *The consistency and cyclicity conditions on R.* As noted in Section 2.2.4, consistency is more than a loose verbalisation (like "the same topic relation has the same name and a different topic

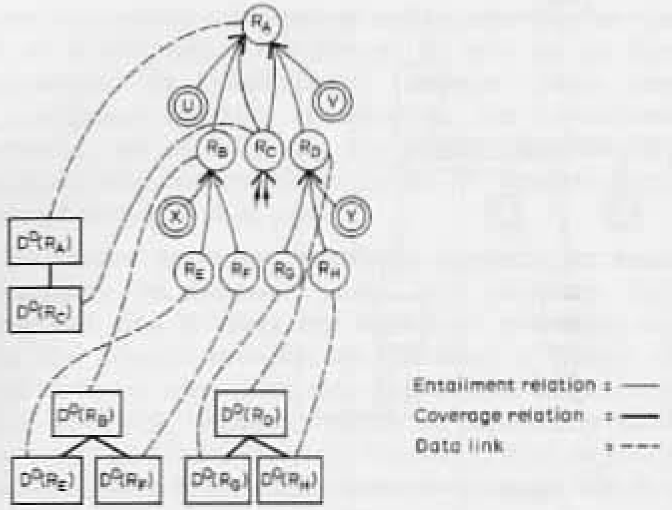


Fig. 7.4 (b). A typical relation between the entailment structure $D^1(R)$ or ES(R) and the task structures $D^0(R_i)$ or $TS(R_i)$. The task structures, represented as command graphs, may be ordered into an hierarchy under command graph coverage but the ordering is usually only local since the entailment structure contains disjunctive substructures (as shown) and analogy relations. The fields of $D^0(R_E)$, $D^0(R_F)$ (as shown above) may not be directly comparable with the fields of $D^0(R_G)$, $D^0(R_H)$ and the pertinent models/demonstrations may or may not be open to execution in the same modelling facility.

relation a different name"). It means that any node named as i, which is pointed out as distinct in $D^1(R)$, has a unique, though generally many faceted, description $D^1(R_i)$ and a unique command graph $D^0(R_i)$. The bite behind this requirement appears at the point when $D^1(R)$ is embodied, at an interface, in ES(R) and the $D^0(R_i)$ are embodied, at the same interface, in the $TS(R_i)$; in which case, these inscriptions must satisfy Fig. 4 (a).

In this connection, it is interesting to notice that the theory is uncommitted in respect to reductionism or philosophical dogmas of the same clan. Knowledge as it is exhibited to the student is not "reducible" and we neither affirm nor deny that it might become "reducible" if the unzip operation of Fig. 2 were extended far enough. The point is made in Fig. 4 (b) and Fig. 4 (c). If the knowable universe of modelling (in case the $D^0(R_i)$ are task structures, the modelling facility) is reducible to homogeneous elements, then all of the $D^0(R_i)$ are hierarchically ordered (for

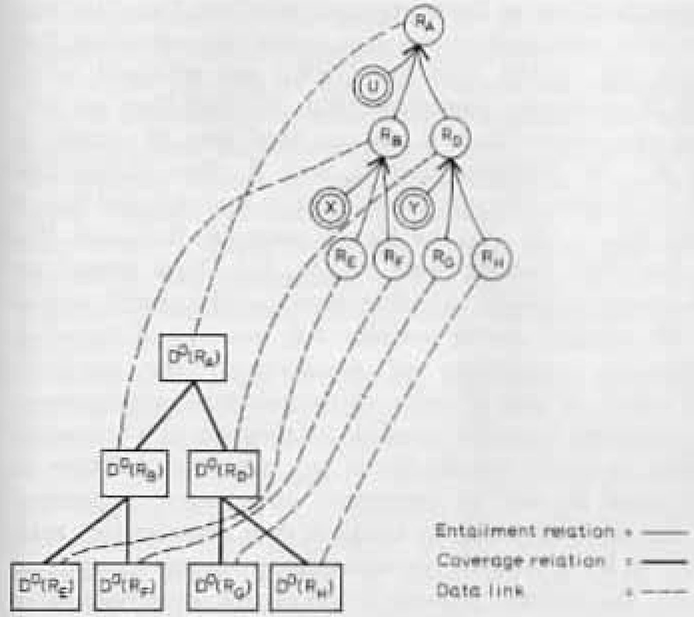


Fig. 7.4 (c). If the entailment structure is locally conjunctive the task structures may be hierarchically ordered under coverage. Such configurations are unusual and the coverage relations are thus suppressed as in Fig. 4(a) to avoid any commitment in this matter.

example, R_i in Fig. 4 (b) establishes a relation between R_j and R_k: R_l establishes some relation between R_m, and R_p; if R_j, R_k, R_m, R_p really are elementary, not merely regarded as elementary, then there is an R_o that heads the lot, and. as corollary, all of these relations can be modelling in a uniform modelling facility). Conversely (Fig. 4c) if there are distinct kinds of entity (ontological classes, perhaps) then there is no reason why R_j, R_k, R_m, R_p, should be modelled in the same facility, nor any reason why $D^0(R_j)$, $D^0(R_k)$, $D^0(R_m)$, $D^0(R_p)$ should be arranged by an hierarchy.

In any case, very few conversational domains are presented as homogeneous in the requisite sense (regardless of whether they might, at some finer grain of exposition, become so). Consequently, there are usually several quite distinct universes in which L^0 explanations are interpreted, and (for non-verbal explanation) several modelling facilities are required. Since there may be as many universes/facilities as there are topic relations Fig. 4 (a), our paradigm, avoids commitment on this score by eschewing the suggestion (of Fig. 4 (b) that the $D^0(R_i)$ are arranged in an hierarchy, or the converse suggestion (Fig. 4c) that they are not. In the special case where there is but one legal way of coming to know (say) R_o, by knowing first R_i and R_j (the conjunctive substructure of chapter 4 reintroduced later in this chapter) then a hierarchy like Fig. 4 (b) may be locally asserted. But even this special case has little bearing upon reality; the "legal ways" are "legal to the source's thesis". Another source might permit several ways where the original source permits only one way of knowing.

The companion condition of consistency, the cyclicity condition of rule 5, is also of some epistemological consequence. It implies a comment upon the possible localisation of a knowable and the extent to which knowledge of any one topic relation R_i (idea, thesis) must be seen as dependent upon some, or perhaps upon all, of the R_i in R. True, in terms of data organisation, each node i (or R_i in R) is written permanently in storage (as part of ES(R)). Hence, the abstract representation of this topic relation (a node in the graph) can be associated with an identity operator that carries it into itself. But this comment refers to a writing or an inscription of R_i: not to the operational form of R_i as a concept (alias procedure) which is part of knowledge. In general, a concept does not have a local identity. The cyclicity condition guarantees (under specific circumstances) that any concept of R_i may be reconstructed in the context of R, by a reproductive process. But this involves at least one other R_j in R[18].

3.5. *Meaning of R_i*. The source's meaning of a topic relation R_i is its reconstruction from other relations in R (as prescribed by a part of $D^1(R)$) and its explanation or realisation as prescribed by $D^0(R_i)$.

[18] Any directed graph in which each node i has an identity Id (where $Id_i(i) = i$) and in which the labelled and directed arcs stand for associative operations (namely, $e \circ (f \circ g) = (e \circ f) \circ g$) is a category. The nodes may represent abstract objects such as sets or functions or relations or complete mathematical systems such as groups or other categories. The operations employed in the present case are associative and the graph of a relational network or an entailment network is a category. Moreover, cyclicity ensures that the graph of the entailment mesh is a "quasi category" or a category with "distributed identity". (A graph with nodes corresponding to each substructure that might be counted as a head is a category unqualified).

For student A, the meaning of R_i is its reconstruction by $\underline{Proc}^1_A$ i in π^1_A and its explanation $\underline{Proc}^0_A$ i in π^0_A.

In either or both of these two senses, R_i's meaning to a source may be greater than R_i's meaning to a student (for example, if A is a novice, there may be no $\underline{Proc}^1_A$ i in π^1_A and/or $\underline{Proc}^0_A$ i in π^0_A, or the existing classes of procedures may be more restricted than those "legally" permitted by the source). On the other hand, R_i may have a "larger" meaning to a student than it does to the source (A can reconstruct and/or explain R_i in ways that are capable of execution, though they are not seen as legal by the source).

These two senses of meaning correspond, rather obliquely, to the classically distinguished categories of connotative and denotive meaning. Surely a reconstruction, either by the source or the student, is one connotation (to the source or to the student). But the explanations of R_i prescribed by $D^0(R_i)$ and $\underline{Proc}^0_A$ i are also connotations. The denotation of R_i becomes apparent only in respect of some universe of elements that may be regarded as elements of a set. Given this auxiliary specification the denotation of R_i is in the classes of elements picked out by the source when he unzips R_i and picked out by a student, A, if he executes $\underline{Proc}^0_A$ i in that context. For example, given a universe of cards showing pictures of Clobbits (the imaginary animals on which the test of chapter 3 is based) the source's (or test designer's) denotation of the relation (between behaviour, physical characteristics, etc.) called "Clobbit", is all the cards. The student's denotation is all, or some of them; those he believes to exhibit the relation in question.

4. Relational operators

Since relations in extension are subsets of product sets it is possible to employ set theoretic operations (Ashby, 1967 or the Bourbaki School) as relational operators. The difficulty is that some set theoretic operations, notably union, are prone to yield trivial results.

We have relied, chiefly, upon a set of specially tailored relational operators, due to Codd (1970), which have several virtues. First, their applications to one or more relations is guaranteed to yield a non-trivial relation as a result. Next, the field of relations in extension is closed under this set of operations i.e.

any relation can be derived from a collection of other relations by an appropriate sequence of operator applications. This set, called Rel op Set is actually redundant. Some relational operations can be derived from others. Moreover, it is often convenient to use rather more relational operators than necessary (for example, analogy, which is an addition to Codd's set).

Finally, since our product sets are generated rather than given it is necessary to augment Codd's set in various ways which are described in Appendix J. It is also of some advantage to introduce the idea that the relation, in extension, is "fuzzy" (in the sense of Zadeh 1968, 1971; Tamura et al. 1971 and Shimura 1973).

Generally it is possible to divide relational operators into two distinct categories namely those that do and those that do not loose specificity. For example, projection (the same as projecting points in a plane, say, onto one co-ordinate) does loose specificity and so do restrictions of one relation by another. On the other hand, permutations and certain types of "join" (or knitting together by special forms of composition) do not loose specificity.

In order to reproduce a relation, its specificity must be preserved. But the fact that specificity is lost locally in a relational network does not necessarily imply that it is lost globally. For example, the loss due to local restriction can be repaired by adjoining information about the co-ordinates of the restriction in an other-than-local manner. An algorithm operating upon a relational network, in particular the algorithm which checks for the existence of cyclicity under any selected head node, must therefore keep track of all points at which specificity is lost. Before the network is approved as cyclic, it must check back to make certain that any specificity (essential to reconstruct the topic relation of the head node), has been replaced at some other point. An algorithm of this type is shown as a program listing in Appendix K and this course assembly program is part of the CASTE course assembly heuristic.

5. Transition from a Relational Network to an Entailment Mesh

The IAs representation, Ω (and later, of R) is a relational network comparable to the structures employed in "artificial intelligence" studies. This is reduced and simplified as indicated in Chapter 4. The operations needed to reduce it to a pruned, cyclic, entailment mesh, are, in principle at least, open to mechanisation.

As a preface, many of the conditions imposed by executing these operations are believed to be unduly strict. I am not primarily a mathematician and do not claim to present elegant procedures. The motivation for each step is briefly outlined at the beginning of the section where it is described and though these steps satisfy the requirements that are stipulated, the exclusions they produce are probably far too stringent. This matter is in hand. The mathematicians in the group have already extended the class of admissible structures and have introduced some radical innovations (none of which contradict the special cases we examine). This work is noted but not exhibited in the present discussion.

5.1 *Relational networks.* A relational network is a finite directed graph with nodes that stand for the L* image of topic relations cited by a source and directed arcs that represent paths whereby one relation is said to be derivable from another (one or

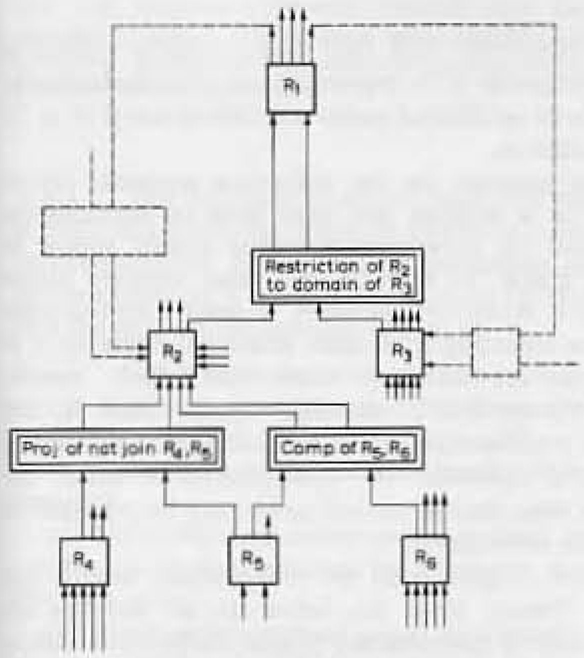


Fig. 7.5. Fragment of a relational network. Labels on arcs indicate relational operators corresponding to derivation of indexed relations. All arcs irrelevant to discussion are deleted but customary richness of connection is indicated by incomplete arcs and (for cyclic paths) dotted lines.

more) relations. Each node in the graph is associated with one and only one label (its name) and with one data link or pointer (drawn as a dotted line to avoid confusion with the arcs of the graph itself) which places the node of R_i in correspondence with one and only one description $D^0(R_i)$ (of how to model, explain, bring about R_i) or, equivalently, the embodiment of $D^0(R_i)$ at an interface (the task structure, $TS(R_i)$), as shown in Fig. 4.

One fragment of a relational network is shown in Fig. 5. Each node has incoming arcs and outgoing arcs. Any cluster of incoming arcs (there may be several for a given node) is covered by a label and signifies a derivation path: for example, R_1 is derivable from R_2 and R_3: R_2 is derivable either from R_4 and R_5 or from R_5 and R_6; the label on a cluster bears a description of the relational operators mustered in making this derivation.

In Fig. 5 the majority of outgoing arcs have been deleted for the sake of clarity. The graph is still unwidely and an anastasmotic plexus of this kind is quite typical, even of well structured subject matter.

5.2 *Entailment networks.* The first step towards obtaining something more manageable is to replace the arc bundle labels by additional nodes. These additional nodes are distinguished (Fig. 5) by a double circle notation.

The distinction is justified on the following grounds: (a) A relational operator is a relation and may thus be legitimately represented as a node. (b) Relational operators usually appear in many places in the graph. To keep the rule, that ordinary nodes have unique names, it would be necessary to draw a confusingly large number of arcs emerging from each relational operator if it were represented as an ordinary node (for, their names, "composition" or "projection", can only appear once in the graph). To avoid a proliferation of uniformative arcs the nodes representing relational operators are distinguished to admit the convention that the same double circled nodes may be repeated at different places in the network.

This transformation, carried out for convenience, should not destroy specificity. Hence, when the labels on arc bundles are converted into clusters of distinguished nodes, distinct derivation paths are preserved by the arrow head notation in Fig. 6. Any ordinary node may have several ingoing arc bundles, just as it may have any number of outgoing arcs. In contrast, distinguished nodes depicting relational operators are devoid of incoming arcs.

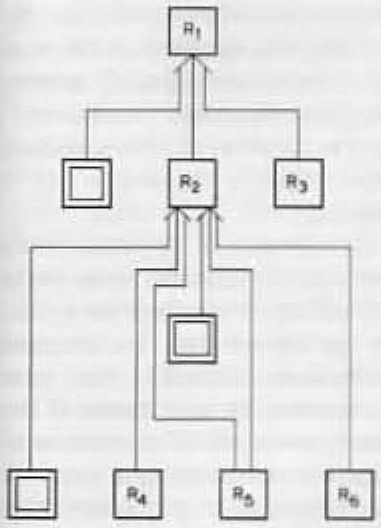


Fig. 7.6. Entailment network obtained from the fragmentary relational network of Fig. 5. Distinguished nodes (shown as double squares) stand for relational operators or any collection of relational operators. The arrow heads indicate distinct derivations.

5.2.1. The structure obtained when the labels are removed from a relational network, when its relational operators are represented as distinguished nodes, and its derivation paths are indicated by arrow notations is called an entailment network. For example, Fig. 6 is the entailment network of Fig. 5.

5.2.2. Let $\Leftrightarrow$ represent a one to one correspondence. Since

(a) Nodes in a relational network $\Leftrightarrow$ Nodes in its entailment network.

(b) Pointers in a relational network $\Leftrightarrow$ Pointers in its entailment network.

(c) Node (topic relations) names $\Leftrightarrow$ Node (topic relation) names.

(d) Arc clusters covered by a label $\Leftrightarrow$ Arrow head notation at the same nodes.

(e) Relational operators noted in a label $\Leftrightarrow$ Distinguished nodes attached, at the arrow head, to the same node.

It follows that a relational network and its entailment network are isomorphic; that is

(f) Relational network $\Leftrightarrow$ Entailment network

5.3 *Discovery and the learning process (entailment meshes)*. In general psychology it is customary to elide the distinction between the kinds of cognitive process that are component parts of learning by grouping these kinds under one general rubric "discovery". There is ample evidence to attest the practicality of this expedient (though, at a theoretical level the implicit suggestion of a "discovering faculty" is a little disquieting).

A comparable generalisation is performed as follows; when transforming a relational network into an entailment mesh (with the entailment network as an intermediary) we subscribe to the principle that mental constructive processes can be grouped together and imaged without distinction. Literally, the term entailment means a permission for discovery by any means at the student's disposal. On the other hand, the kinds of process have been characterised up to isomorphism in the relational network and the use of an entailment mesh as a permission giving structure (a picture of what may be known) is legitimate if and only if, for the student A in question, there are $\underline{\mathrm{Proc}}^1_A$ in π^1_A isomorph to all the relational operators that feature in the corresponding relational network (Chapter 4). These are the L^1 primitive operators, $\underline{\mathrm{Prim}}^1$ of Chapter 4. Because they are assumed to exist, the distinguished double nodes in Fig. 4 (representing relational operators) are left without data links (dotted connecting lines). Otherwise, these operator nodes would be associated with procedures, $\underline{\mathrm{Prim}}^1$, stated explicitly as fuzzy L^1 algorithms.

Surely, an entailment mesh can be drawn willy nilly, whether this statement is true or not. But its adequacy as a permission giving structure is empirically determined for each individual. Strictly speaking, an answer to the question "Do $\underline{\mathrm{Proc}}^1_A$ isomorph to each relational operator exist in π^1_A" ought to be determined by any operating procedure founded on the entailment mesh. All of the CASTE conversational B-heuristics do contain appropriate tests. But even without this facility, the entailment mesh is a cogent tutorial instrument just because the question is often (or in our experience of teaching and learning always) answered affirmatively. On this account, the use of entailment meshes can be justified without specific empirical testing. Under these relatively loose conditions, there is a probability, not a guarantee, that the permissions given by the entailment mesh can be realised.

5.3.1. To summarise, the relational network says what concepts

$\underline{\mathrm{Proc}}^0_A$i may be learned if $\underline{\mathrm{Proc}}^1_A$ isomorphic to the relational operators (with intensions or procedures $\underline{\mathrm{Prim}}^1$) exist in π^1_A and if certain $\underline{\mathrm{Proc}}^0_A$i also exist in π^0_A. Further, it says how they may be learned i.e. what cognitive operations should be applied in order to realise each derivation. Like a relational network, the entailment mesh says what concepts may be learned (given the same preconditions). It does indicate separate derivation paths and thus, for example, the number of different and permissible ways of learning a concept. But, in contrast to the relational network, the computations needed to realise a derivation are deliberately obscured. "Entailment" gives permission for "discovery"; using any kind of computation at the student's disposal.

Though the entailment mesh on its own, furnishes less information than the relational network it is a more readily comprehended and manageable instrument. However, there are circumstances under which the potential information "lost" in transforming the relational network into an entailment mesh is required. For example, since there are usually many derivation paths, the student may require more detail in order to select one of them; similarly the teacher should (according to the tutorial B-heuristic) be in a position to evaluate the student's selections or to recommend a particular class of derivation paths. For either purpose, the "lost" data must be retrievable and consequently it is stored in an operator data base indexed in register with the nodes of the entailment mesh and conditionally accessible, throughout an A, B, conversation, to both participants.

5.3.2. The operator data base is organised as follows. The ith location corresponds to the ith node in the entailment mesh and provides separate storage capacity for each derivation path leading to the ith node. In each compartment of the ith location is inscribed the names of the relational operators (specified in the relational network) that have the ith node as their range (under that derivation) together with the names of their domains i.e. an ordered set of all the distinguished nodes in the entailment network; belonging to this derivation path and deleted in forming the entailment mesh.

5.3.3. Let "$\underset{E}{\Leftrightarrow}$" represent a homomorphism, preserving the relation "entailment" and "$\Leftrightarrow$" (as before) represent an isomorphism,

Relational network $\Leftrightarrow$ Entailment network (Section 5.2)
Entailment network $\underset{E}{\Leftrightarrow}$ Entailment mesh
Entailment network $\Leftrightarrow$ Entailment mesh $\times$ Operator data base
(from Section 5.3.2)

5.4. *Primitive nodes.* Nodes in the entailment mesh that are devoid of incoming arcs are called primitive. The operational significance of this appelation is as follows: if the mesh is used as part of the entailment structure D^1 (R) for a tutorial conversation, then the student, A, must treat the relations they designate as properties he can evaluate.

The requisite condition is dually stated;

"The Student A has a reproducible concept for the relation designated by each primitive node".

Or

"There exist in (π^1_A, π^0_A) a pair $(\underline{\mathrm{Proc}}^1_A\,i, \underline{\mathrm{Proc}}^0_A\,i)$ for each node i, that is called primitive".

There is no obvious way of determining whether or not this condition is satisfied before a conversation starts, but the issue is empirically determined by a tutorial B-heuristic which either does or does not mark all primitive nodes as understood. If not, the conversation is discontinued, at any rate until the student has undergone appropriate training.

Nor is there any way of telling how many nodes, over and above those called primitive, will be marked as understood. This is irrelevant to the conduct of a tutorial conversation, unless the converse condition is obtained i.e. all of the nodes are marked as understood. If so, the student is so knowledgeable that there is nothing he can learn about the domain in question and conversation is also disallowed.

5.5 *Distinguished nodes.* Since distinguished nodes (standing for relational operators) and primitive nodes both call for the satisfaction of certain empirically determined initial conditions they are grouped together as initial nodes.

Thus,

Initial nodes = (Distinguished nodes, Primitive nodes)

6. The form of Entailment Meshes

An entailment mesh is usually a complicated structure since outgoing arcs from any one node re-enter many other nodes as

incoming arcs. In order to extract the organisation of the mesh it is convenient to isolate fragments of it in the imagination as substructures from which most of the outgoing arcs have been stripped away (not in reality, only in the pictures used to describe the mesh).

This trick is particularly useful in establishing substructure definitions that are required either to test an entailment mesh for cyclicity, or to formulate the CASTE heuristics (Chapter 4).

6.1. *Conjunctive substructures.* Select any non primitive node, α. It stands for a relation $R\alpha$, and is at least associated with the substructure shown in Fig. 7.

This picture is verbally construed as

"$R\alpha$ is a *system* of $R\beta$ and of $R\gamma$"

or

"$R\alpha$ *entails* $R\beta$ and $R\gamma$"

or

"Given $\underline{\mathrm{Proc}}^0_A\,\beta$ (a concept for $R\beta$) and $\underline{\mathrm{Proc}}^0_A\,\gamma$ (a concept for $R\gamma$) any A who has $\underline{\mathrm{Proc}}^1_A$ in π^1_A for each distinguished node may construct $\mathrm{Proc}^0_A\,\alpha$ (a concept for $R\alpha$)".

6.2. *Disjunctive substructures.* Two essentially different organisations exist in the majority of entailment networks; one of

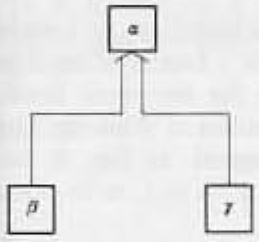


Fig. 7.7. Conjunctive organisation in the entailment network.

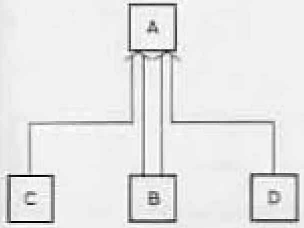


Fig. 7.8. Disjunctive organisation in the entailment network.

them conjunctive in form and one of them disjunctive. For the conjunctive form (Fig. 7) relation $R\alpha$ entails both relations $R\beta$ and $R\gamma$, but no other entailment is permitted. For the disjunctive form (Fig. 8) relation A entails either relations B and C or relations B and D or both.

6.3 *Immediate entailment sets.* Consider the set of all incoming arcs to a given node i. These arcs originate at nodes which are one stage removed from node i. The set of incoming arcs is called the (unqualified) immediate entailment set of node i and is abbreviated to Im Ent Set i. For example, in Fig. 7, Im Ent Set α is (β, γ).

If the organisation of arcs impinging upon a node is disjunctive (as it usually is, in practice) several Im Ent Sets exist, one for each conjective cluster of incoming arcs. These sets are indexed by an integer, k. For example, in Fig. 4, where R = 1 or 2; Im Ent Set 1, A = (B, C) and Im Ent Set 2, A = (B, D). The unqualified Im Ent Set i is the union of these k indexed sets; thus, Im Ent Set A = (B, C, D). On interpretation, nodes in Im Ent Set k, i, stand for the relations which must be cited in the kth type explanation of how to know of R_i, and for which procedures (alias concepts) must be available if R_i is to be learned or constructed. Similarly, the nodes in Im Ent Set i stand for the relations that might be cited in any permissible explanation or reconstruction of R_i.

6.4 *Kernels and entailment sets.* The entailment set of a node in a fully conjunctive network is the union of a family of Im Ent Sets obtained by iterating the construction for the node itself, each member of its Im Ent Set, for each member of these Im Ent Sets, and so on, to a depth d from the original. In Fig. 9, for example, Im Ent Set α of depth d = 2 is $(\beta, \gamma, \delta, \epsilon, \zeta, \eta, \theta)$. This

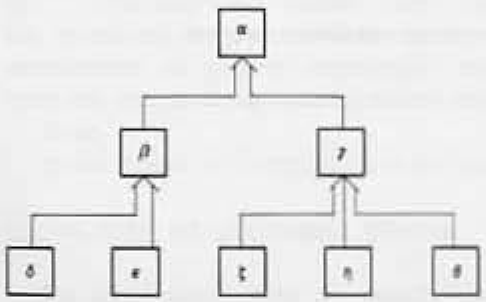


Fig. 7.9. Conjunctive entailment network with unique Ent Set α.

structure is also (because the network is fully conjunctive) called the *kernel* of α to depth d; that is, kernel $\alpha(d)$.

If the network contains nodes with disjunctively organised ingoing arcs, several kernels exist. They are obtained explicitly by selective iteration over fully conjunctive paths. In Fig. 10 for example, there are 3 kernels of A for depth d = 2. They are kernel 1, A = (B, C; F, G, H, I) and kernel 2, A = (B, C; F, G, J, L). Like the distinct Im Ent Set k, i, the kernels of R_i are k indexed as kernel k, i. The union of the kernels at depth d is Ent Set i. Thus Ent Set A = (B, C, D, E, F, G, H, I, J, K, L). Though the value of d is arbitrary in the definition, various criteria are used to fix it with reference to particular data structures.

On interpretation, the nodes in Ent Set k, i stand for those relations that must be associated with procedures (alias concepts) in order to construct or explain R_i in the kth way; the nodes in Ent Set i stand for relations that may be cited in order to give any type of explanation of relation R_i.

6.5. *Chain structures.* An entailment chain is a sequence of arcs, pointing in the same direction, that is outgoing from a node y and ingoing to a different node x. For example, in Fig. 10, if y = A and if x indexes nodes at depth d = 2 there are the following entailment chains of lengths 2: (F, B, A); (G, B, A); (G, C, A); (H, C, A); (I, C, A); (I, D, A); (J, C, A); (R, D, A); (L, E, A); (J, C, A) and (L, E, A). The entailment chains in a full entailment mesh (rather than an imaginary fragment like Fig. 10) may be

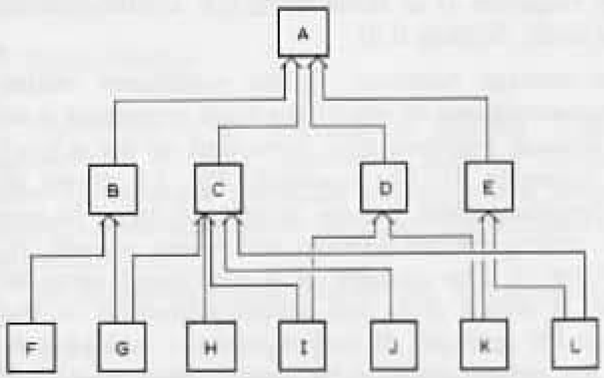


Fig. 7.10. Disjunctively organised entailment network in which there are several distinct Ent Sets A.

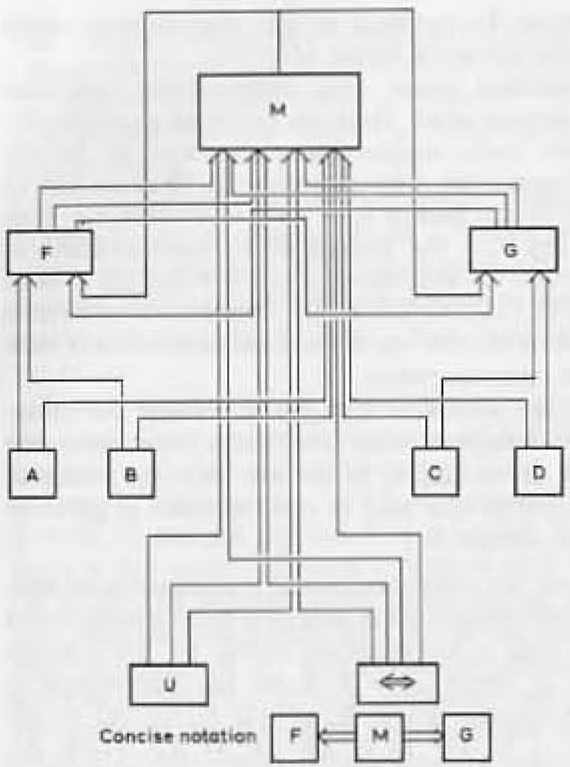


Fig. 7.11. An analogy relation.

cyclic or they may originate at an initial node (i.e. a distinguished node or a primitive node, Section 5.5).

6.6. *Cycles and analogy relations.* Cyclic entailment chains feature in several substructures of which the most important is an *analogy* relation between relations that form part of the subject matter (Anal of Appendix J). For example, Fig. 11 shows an analogy between functions (that is, one to one or many to one relations) F and G, having distinct domains and ranges, namely, A, B for F and C, D for G. The analogy M is contingent upon an analogical universe, $\mathcal{H}$ which picks out salient properties of the domains and ranges; for example, M may represent a metaphor or it may represent the correspondence between a mechanical law (say the functional specification of friction in a universe of blocks and inclined planes) and an electrical law (wired up, in a universe

of integrators, resistances and capacitors, in an analogue computer).

The form of the analogy is

$$(A\,(F)\,B) \underset{\mathcal{H}}{\Leftrightarrow} (C\,(G)\,D)$$

or "α is to β as γ is to δ" (α an instance in A, β an instance in B; γ an instance in C; δ an instance in D; F the relation between A and B, G the relation between C and D. Further F is the isomorph of G, $(F \Leftrightarrow G)$ in the analogical universe of $\mathcal{H}$ (a less stringent condition, namely that F is the homomorph of G in the universe of $\mathcal{H}$ $(F \Rightarrow G)$, is possible but all further comments are confined to the case of $(F \Leftrightarrow G)$.

Finally M is the relation, equivalently stated as a definition of the distinguishing property $\mathcal{H}$ in terms of the relations established between the domain/range pairs of F, G, when $F \Leftrightarrow G$ or as the distinguishing property of $\mathcal{H}$ because of which $F \Leftrightarrow G$ is a valid relation. (Hesse (1963), for example). Hence, given M and F, G is obtainable; given M and G, F is obtainable. The relation M may either be obtained by citing the relations F and G together with $\mathcal{H}$ and $\Leftrightarrow$ or by citing the domains and ranges of F and G together with C, G and $\Leftrightarrow$. Since analogy relations are very common the substructure in Fig. 11 is replaced by the special symbol "$O \Leftarrow O \Rightarrow O$" where the outermost nodes are F and G (in Fig. 11). This is the notation introduced in Chapter 4, Fig. 7.

6.7. *Cyclic submeshes and reproducibility.* Any analogy (under $\Leftrightarrow$) is a cyclic submesh. Moreover (as a consequence of this), any analogy under $\Leftrightarrow$ is reproducible[19].

7. Cyclic Meshes

Cautionary note: Throughout Section 7 and Section 8 the symbols U, V, X, Y, previously reserved for other purposes, are used to designate particular values or sets of values of the node index, i. The moratorium on other uses of these sybols is rescinded in Section 8.8.

An Entailment mesh may or may not be cyclic (Section 6.5 and

[19] Analogies involving homomorphisms may be reproducible also, depending upon whether or not the distinctions lost in the many to one correspondence are restored by adjoined relations.

6.6) depending upon the existence and arrangement of cyclic entailment chains.

7.1. *Motivation for testing the cyclicity of an entailment mesh.* If a consistent entailment mesh is cyclic, it constitutes a viable domain for a P Individual, A in the following sense. If π_A^0 contains $\underline{Proc}_A^0$ r to satisfy or bring about all relations denoted by primitive nodes r in the mesh and π_A^1 contains $\underline{Proc}_A^1$ r and also $\underline{Proc}_A^1$ = $\underline{Prim}^1$ isomorph to the relational operators that are distinguished nodes in the corresponding entailment network (but deleted from the entailment mesh) then A can evolve in (i.e. learn) the domain designated by the mesh. In other words, A can build $\underline{Proc}_A^0$ s (alias concepts) to satisfy all of the relations denoted by nodes in the mesh; and without further requirements A can reproduce all of these concepts as memories.

In a strict conversation, a record showing that concepts for all of the relations have been learned in any of the ways permitted by the mesh is evidence that the conditions of the last paragraph are satisfied. In other words, if the mesh is cyclic and if it has been learned (by constructing $\underline{Proc}_A^0$ i for each R_i) these $\underline{Proc}_A^0$ i s may be reproduced (as memories of the R_i) whilst the P Individual remains in this domain.

7.2. *Testing the cyclicity of an entailment mesh.* A cyclic mesh contains one or more entailment chains. Any chains it does contain are connected (guaranteed, by construction; Section 2) and the mesh is consistent (also guaranteed). Any node in a cyclic mesh either belongs to a cyclic entailment chain or it is primitive.

7.2.1. Cyclic entailment meshes of the kind shown in Fig. 12 are of theoretical interest but of little consequence in the present

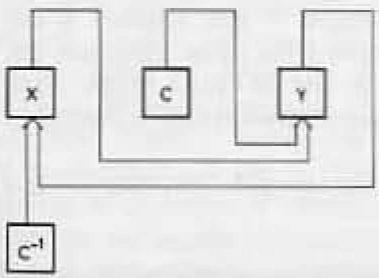


Fig. 7.12. A vacuous cyclic entailment chain, such as an inverse operation (for example, permutation) on any domain.

argument, because there is no "entry point"; they are called vacuous.

7.2.2. Other cyclic meshes are characterised as having head nodes x and tail nodes y according to the way they are drawn. All of them have at least one node (one in Fig. 13, two in Fig. 14) with disjunctively organised incoming arcs. The appelations "head" and "tail" do not depend upon the relations that are designated by these nodes. For uninterpreted meshes, the names

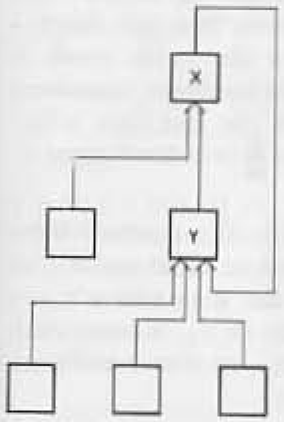


Fig. 7.13. A cyclic entailment chain with disjunctively organised ingoing arcs to node Y.

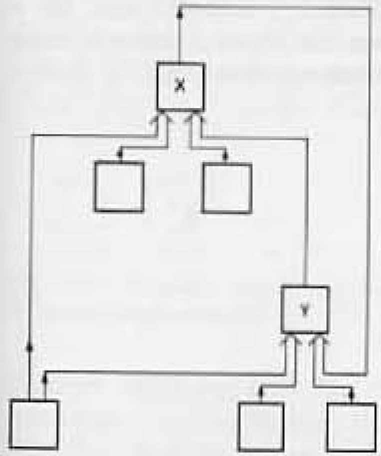


Fig. 7.14. A cyclic entailment network with disjunctively organised ingoing arcs to node X and node Y.

"head" and "tail" refer (as stated) to how it is decided to draw the mesh. For interpreted meshes, these names may reflect a more thoughtful description of the mesh; they still do not depend upon the mesh itself.

7.3. *Head node.* Suppose a node, X, is called the head of an arbitrary mesh. If the mesh is cyclic (and, henceforward, not vacuous) then it satisfies the (following) requirements, 7.3.1, 7.3.2 and 7.3.3.

7.3.1. In each Kernel k, X, there is, at depth d_k at least one node Y_{kj} (in general, there are many of them; thus the index j) such that X is in Im Ent Set Y_{kj}. Since the cyclic mesh is non-vacuous the incoming arcs to Y_{kj} are disjunctively organised; hence, there are several Kernels l Y_k and Im Ent Sets l Y_{kj}; l = 1, ... M. Suppose the outgoing arc from X is in the Kernel l_X Y_{kj} (i.e. that X is in Im Ent Set l_X Y_{kj}).

7.3.2. The other arcs, incoming to Kernel l_X Y_{kj}, emanate from other nodes in Im Ent Set l_X Y_{kj}. These one or more nodes (call them λ) form a class, Λ. If the mesh is cyclic, then either Y_{kj} is simple; all λ in Λ belong to Ent Set k, X, (dk) or Y_{kj} is supported; all λ in Λ belong to all M distinct Ent Set l Y_{kj} or both conditions hold; one for some nodes; one for the others.

7.3.3. Conditions 7.3.1 and 7.3.2 hold for each Ent Set kX in Ent Set X.

7.3.4. On interpretation; if condition 7.3.2 holds true for a simple Y_{kj} then each R_X is a relation for which a concept must have been constructed in reaching a concept satisfying R_X from a

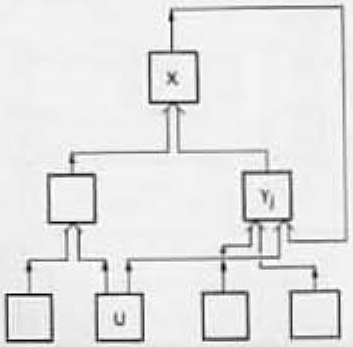


Fig. 7.15. A form of simple closure.

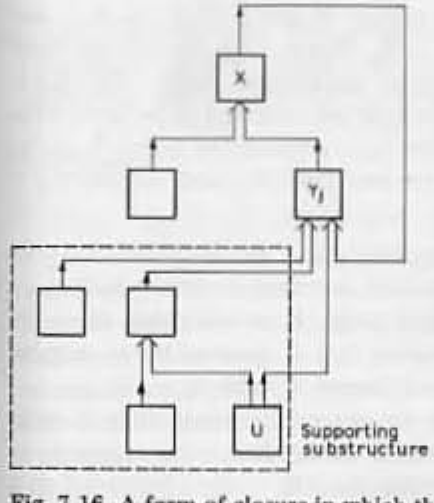


Fig. 7.16. A form of closure in which the status of Y_j as a tail to the head X is maintained by a supporting substructure.

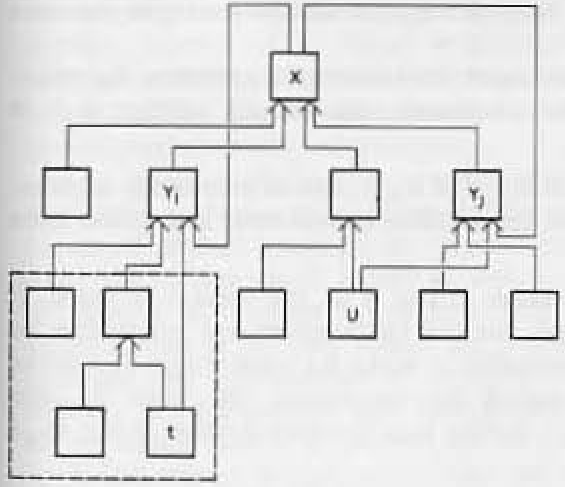


Fig. 7.17. Example of a mixed condition: Y_i has a supporting substructure; Y_j shows simple closure.

concept satisfying each $R_{Y_{kj}}$ (the situation shown in Fig. 15); if condition 7.3.2 holds true for a supported Y_{kj} then a concept for each R_X must have been constructed to obtain a concept for each $R_{Y_{kj}}$, regarded as part relations of R_X in Kernel k, X (Fig. 16). The situation corresponding to the mixed condition is shown in Fig. 17.

7.4. *The case of several head nodes.* If and only if the R_{X_c} are analogous (Section 6.6 or Fig. 10), there may be several head nodes X_c: $c = 1, \ldots Z$ in a cyclic entailment mesh. The same criteria apply apart from: (a) The tails are indexed in register with the heads; thus Z different indices Y_{kcj} replace the index Y_{kj}. (b) All members X_c of the head X are assigned the same depth of d = 1.

7.5. *Cyclic meshes.* The requirements of Section 7.3 (if anything, unduly strict requirements) are used to test whether or not an entailment mesh contains a cyclic entailment submesh under a head node X (or head nodes X_c). In general there may be several cyclic meshes spanning different depths $d_{k1}, d_{k2}, \ldots$ from X. Hence, supposing there is any cyclic mesh, it is usually necessary to select amongst several possibilities by choosing a vector of depths $\Delta = < \delta_k >$ so that the kth cycle is bounded at a given depth (for example, $\delta_k = d_2$).

7.6. *Choice of head node:* An interpretation. At one of the instants prescribed in Section 2 the IA asks the source to choose a tentative head.

On interpretation, the head node stands for a relation R_X which the source believes to adumbrate the subject matter, as it is currently expounded.

7.7. *Tests.* Using this X (or, if R_X is part of an analogy relation, a set (X_c) the IA tests for whether or not there is a cyclic mesh with head X.

7.8. *Acceptance of mesh.* If this is so, the head is accepted. If not, it may be possible for the IA to point out what kinds of explanations would be needed to make the mesh cyclic. If suitable explanations are furnished by the source, the head is again accepted. Failing that a further head is selected either at this stage or later.

7.9. *Pruning for one head node.* If the head is accepted it is the head node of one or more cyclic entailment meshes. Supposing there is only one, this mesh is pruned (as below) and displayed to the source as a graphic representation.

7.10. *Pruning for several head nodes.* If there are several cyclic meshes the source is required to stipulate a vector of depths, Δ which selects one mesh and this is pruned.

7.11. *Depth.* On interpretation; specifying Δ influences how "far back" in the subject matter the source wishes to go i.e. the expected initial "level of accomplishment" of the students-to-be.

8. Pruning

Even very small entailment meshes, for example the mesh in Fig. 15 are fairly complex graphs. Moreover, given an all embracing statement that each relation denoted by the mesh is reproducible, the arcs signifying cyclic parts of the reproductive paths are redundant. Hence these arcs, which often outnumber the others, need not be displayed to a student. They are deleted and the cyclic entailment submesh is severed from the original mesh by a process called pruning to yield a pruned cyclic entailment mesh (Section 1.2) which is the basis for an entailment structure $D^1(R)$ or its physical embodiment in an interface ES(R).

8.1. *Motivation for pruning the entailment mesh.* One reason for pruning the entailment mesh has already been mooted; namely, clarity of a visual representation in $D^1(R)$ or its inscription ES(R). Once the cyclic character of each topic relation is pointed out, the essential structure is more prominent in the pruned mesh than the original mesh.

As a result of pruning, the mesh is rendered quasi-hierarchical. Except for analogies between parts of the subject matter (these are preserved under pruning) the pruned mesh forms a partially ordered structure which, locally at least, is an AND/OR graph (like the graphs of Boolean expressions used to express implicants). Hence, there is another reason for pruning the mesh. Provided analogies are distinguished or recognised (as they may be, very readily) the B-heuristics that handle ES(R) may be written to accommodate the well known search procedures applicable to AND/OR graphs; similarly (but with the same reservations) it is possible to employ standard methods for finding minimal paths which are needed in order to specify a rational teaching strategy.

8.2. *Form of the pruned entailment mesh.* The hierarchical appearance of the pruned mesh is misleading insofar as it (deliberately) obscures the cyclic relations between topic relations. Hence, the following salient points about pruned entailment meshes deserve repetition.

8.2.1. Each relation is reproducible, even though (chiefly for graphical convenience and readability) the entailment chains representing permissible reproductive processes have been deleted.

8.2.2. Although the description looks like an hierarchy and thus a taxonomy of nouns, the knowables represented are inter-linked relations; knowables form part of a "verb net" not a "noun tree".

8.2.3. The choice of X (and possibly Y) is made at the discretion of a source. Very likely it reflects his experience as a teacher. But, whatever else, it has no direct bearing on the relations depicted in the entailment mesh.

8.2.4. However, the choice of X does impose a description on the mesh. In selecting X the source is forced to say (at least) that X is superordinate to some Y and that any other node is subordinate to X. The "superordinate/subordinate" description is obtrusive in the pruned mesh; roughly as its "top and bottom".

8.2.5. That this is so does not imply an immutable ordering on the relations R_X and R_Y. For example, the notion of an "operon" can lead to "repressor" and "operator gene" (the relations in molecular biology cited in Section 2.1.5.1) or vice versa.

8.2.6. In choosing a head node like X, the source commits himself to one polarity or the other; it is up to his discretion how he chooses to describe the related notions. Moreover, one source may choose one head, another source another head; and the same source may choose different heads with different student populations in mind.

8.3. *Pruning operations.* The cyclic entailment mesh is pruned by carrying out the following instructions.

(a) If there is only one head node, X, go to instruction (b); otherwise go to (f)

(b) Delete all outgoing arcs from X, listing those that terminate on the ingoing arrows of nodes in Ent Set X to identify the loci of the Kernels $l_X Y_{kj}$.

(c) For $\delta_k \geqslant d_k$ and for each k in turn delete all outgoing arcs that are not incoming either to X or to some node in Ent Set X.

(d) For each tail Y_{kj} in Ent Set k, X, (d_k), for $\delta_k \geqslant d_k$.

If Y_{kj} is a simple tail node (Fig. 15), delete all incoming arcs to all Im Ent Sets l Y_{kl} for $l \neq l_k$ calling the relations denoted by these nodes primitive and annotating the nodes.

Otherwise, if Y_{kj} is a supported tail node (Fig. 16) delete all incoming arcs to nodes in any Im Ent Set (support of Y_{kj}) calling the relations denoted by these nodes primitive and annotating the nodes.

(e) If a submesh headed by node X is severed from the original entailment mesh the pruning is complete; otherwise sever any entailment chains entering nodes in Ent Set X, below the Im Ent Set of the node in question, calling the relations designated by these nodes primitive and annotating these nodes.

(f) Replace all occurrences of "X" in instructions (c), (d) and (e) by "in any X_e"; of "Y_{kj}" by "Y_{kej}" and of "$l_X Y_{kj}$" by "$l_{X_e} Y_{kej}$" and go to instruction (b).

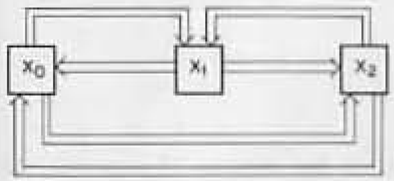


Fig. 7.18. Set of analogous head nodes.

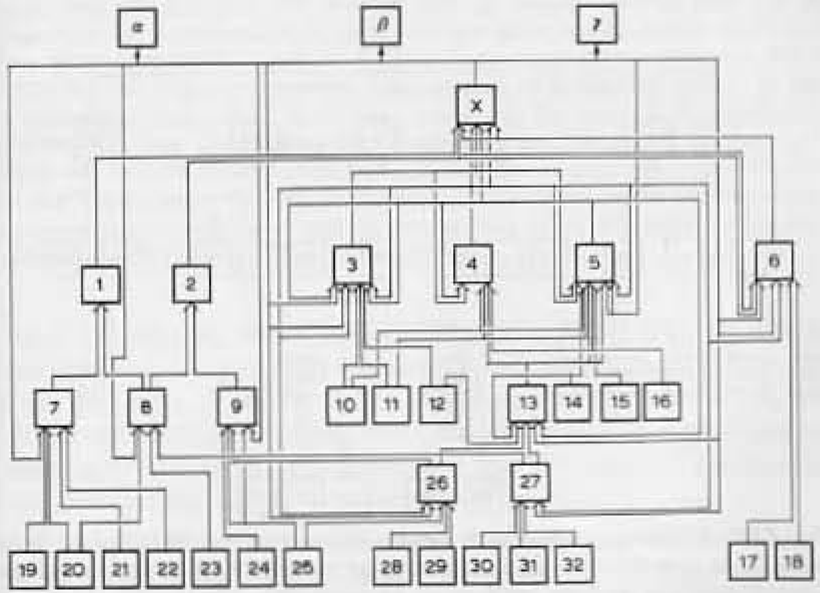


Fig. 7.19. A cyclic entailment mesh: notice the analogy relation R4 between R3 and R5. Only some of the connections that are deleted by the rule of Section 8.4 are shown in the figure.

8.4. *Deletion*. Following this, redundant connections (if any) within the pruned entailment mesh are deleted by the following procedure.

Consider all configurations of the type shown in Fig. 20. If it happens that U can be constructed from Y (that is, U is formed so that Y is devoid of operations like <u>Proj</u> and <u>Rest</u> which loose specificity then any arc linking Y and X in Fig. 20 is deleted. The

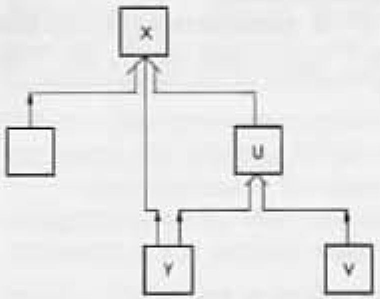


Fig. 7.20. Deletion of redundant arcs. If the relation U preserves the specificity of V and Y (which is transmitted without loss to X) then the connection linking Y to X is deleted.

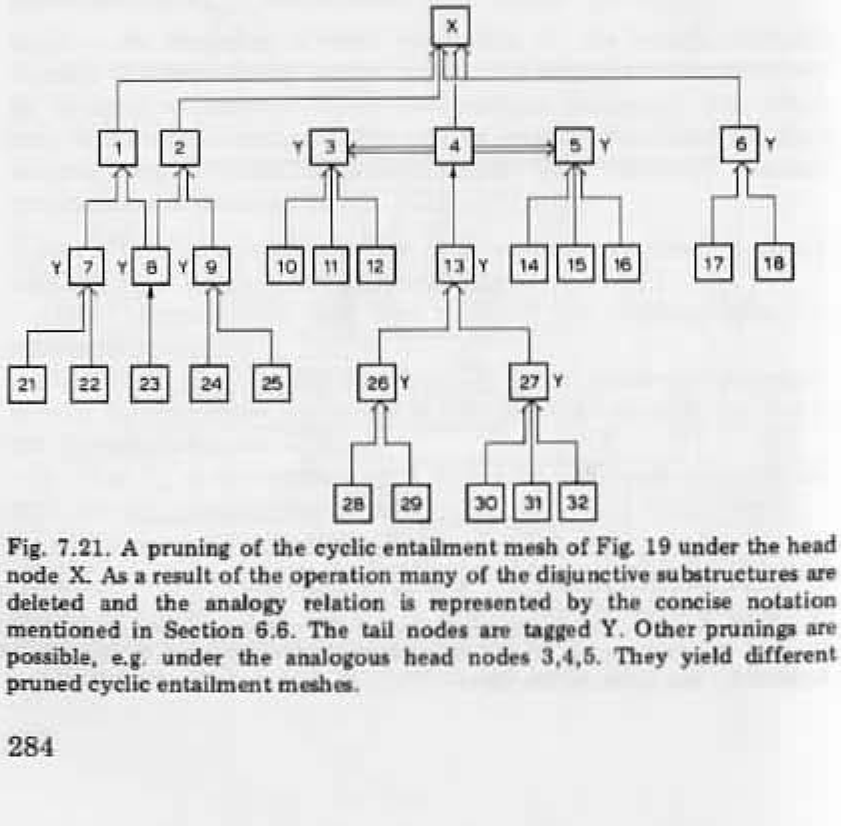


Fig. 7.21. A pruning of the cyclic entailment mesh of Fig. 19 under the head node X. As a result of the operation many of the disjunctive substructures are deleted and the analogy relation is represented by the concise notation mentioned in Section 6.6. The tail nodes are tagged Y. Other prunings are possible, e.g. under the analogous head nodes 3,4,5. They yield different pruned cyclic entailment meshes.

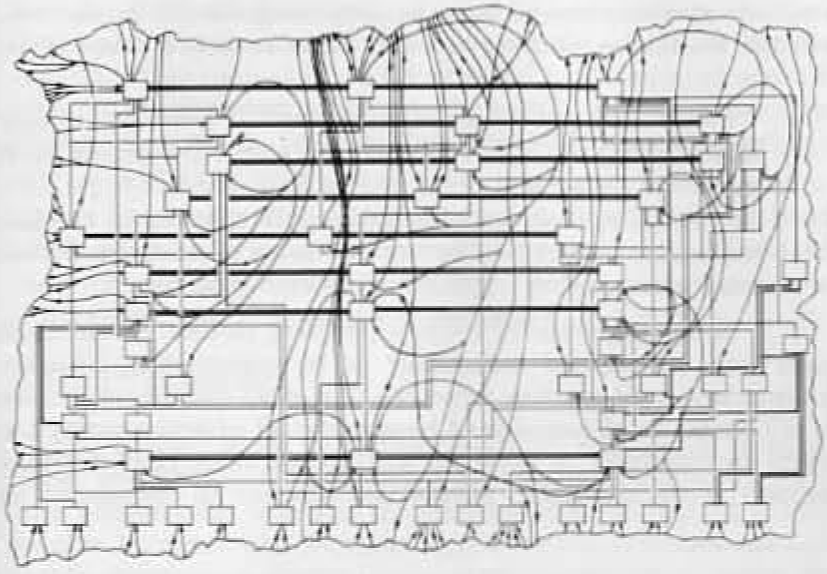


Plate 9. Part of unpruned cyclic entailment mesh from which the entailment structure of probability theory is derived. For clarity the cyclic components have been redrawn over the pruned mesh as irregular curving lines and the ragged border is penetrated by arcs to other nodes within the boundaries of the (pruned) statistics mesh in Fig. 10. Some cautionary comments are in order (a) This does not represent the structure of probability theory: at best it represents a thesis about this subject matter (b) This thesis is not completely represented. Any competent source is able to invent and justify far more connections. But the relative density of connections is genuinely representative of the source thesis. Notice, for example, the high density of analogical connections in the middle part and the connections at the left, representing analogies with many kinds of real life experiment.

result of pruning the cyclic entailment mesh in Fig. 19 and of carrying out the prescribed deletions and completions is shown in Fig. 21. Plate 9 is the cyclic entailment mesh from which the entailment structure of the main example (Fig. 3) is obtained by pruning and deletion. It indicates the real life magnitude of meshes and justifies the EXTEND program of Appendix K.

8.5. *Limitation*. As a convention, R (stated in L*) is the relation in the cyclic entailment network restricted to just those relations designated by nodes in the pruned cyclic entailment mesh (so for example, R does encompass the cyclic relations). Using the L* index, i, number the nodes in the pruned cyclic entailment mesh uniquely and thus the relations, R_i, they designate. All R_i are in R, though R contains relations over and above those in a union of the R_i.

8.6. *Status of a pruned entailment mesh*[20]. I can see no easy way of describing the status of a pruned cyclic entailment mesh with respect to the general cyclic meshes from which it might have been derived. The basic and intriguing difficulty lies in the fact that head and tail are chosen before the extent of the pruned mesh is known, or even before a cyclic mesh is known to exist.

8.7. *Some properties of cyclic entailment meshes.* Several facts have been established empirically or by constructing plausible extrapolations of existing networks. For most entailment meshes it is possible to choose several head nodes all of which satisfy the cyclicity condition. If accepted and pruned these yield different pruned cyclic entailment meshes and (Section 9) correspond to

[20] At the time of writing this manuscript I did not recognise the broader implications of the fact that all relations between topic relations are morphisms of some kind or the full implications of the fact that isomorphic analogies have a specially reserved status. Hence, these points are not stressed and the statement of Section 8.6 quite accurately reflects our thinking at that time. During the last year, however, a more thorough appreciation of what is involved has led to a relatively sophisticated theory in which pruned cyclic entailment meshes can be characterised. We are currently in a position to comment upon the semantic interpretation of meshes and the way that a particular mesh is embedded in a body of knowledge much wider in scope than the source thesis. These developments are reported in Volume II.

One clue to the developments lies in the rule in Section 7.4 at first sight rather arbitrary, that assigns isomorphic analogies to the same level under the "superordinate/subordinate" descriptor, on choice of a head or heads. The other clue is the recognition that the relational operators are a method, as it turns out a very effective method, for externalising the syntactic constraints in a thesis directionally; but that, all the same, the connections are morphisms between topic relations from which isomorphisms are singled out for attention. The terms of (or relations related by) an isomorphism may or may not (but, lacking further evidence, must be assumed to) have task structures or models interpreted in distinct universes of interpretation. The semantic distinctions that do exist are obtained by eliciting descriptors (in the manner described in later sections) over the terms of all isomorphic analogies. For this purpose the isomorphic analogies must be elicited with priority and assigned (as they are at the moment, though without special justification) to distinct levels or values of the "superordinate/subordinate" descriptor. Hence, the entailment mesh consists in layers of potential isomorphic analogies (some, perhaps, not realised but all of them with terms in the same layer).

basically different descriptions of the same underlying relations. The point is not vacuous, for it is easy to construct counter-examples that have only one acceptable head node. In practice, such structures appear to be rare. Something is known about the variety of possible prunings for different types of entailment mesh, but this work is still in progress.

8.8. *Cautionary note.* Henceforward, the symbols U, V, X, Y, which temporarily stood for particular values of i are reserved for their former and normal purpose; X the generic term for an input; Y the generic term for an output and so on.

9. The L^0 Description

Any pruned cyclic entailment mesh, R, has certain primitive nodes which the source stipulated in choosing the depth vector Δ. Any student who learns in the conversational domain of R must be able to regard the relations tagged by the primitive nodes as simple properties. Hence, the names of primitive nodes are predicates in L^0 and the descriptions in $D^0(R_i)$ are L^0 expressions involving these predicates.

9.1. If the L^0 explanations permitted by $D^0(R_i)$ are to be elicited non verbally, in a modelling facility, then this facility must incorporate objects imbued with the properties that are predicated.

For example, STATLAB was designed on the basis of primitive nodes in the probability theory entailment structure (the lower-most nodes in Plate 3 of Chapter 4). The STATLAB design process (which is typical of any modelling facility) has two parts.

(a) To specify artifacts having the properties of the primitive nodes because they satisfy the relations stated on the labels of these nodes (so, for example there are sockets/signal lamps with the property "simple events"; there are operations with the property "+" and "–", that is, operations to realise the triadic relations "+" and "–" since these appear as labels on the primitive nodes).

(b) To specify enough artifacts of a given type to guarantee that each correct explanation can be complete and correct (so, for example, there are 8 simple event sockets; the minimum needed to model the exclusive disjunction of composite events of which, for generality, there must be more than two, and to do so in the context of a modelling universe that may have less than 8 members. Any set of 7 sockets will suffice, one being deliberately excluded from the picture).

9.2. This progression, from $D^1(R)$ and its primitive nodes to the design of a modelling facility is typical, though conceivably it could be reversed (i.e. the modelling facility might be in the source's mind before he expands his theory of the subject matter). In either case, it is necessarily true that any modelling facility is described by the L^1 description in $D^1(R)$: for example, the quadrants of STATLAB are partitioned, as the entailment structure is partitioned, by the descriptions "structural model/measure" and "real experiments/abstract model or measure" (in the notation of chapter 4: the distinctions "C or G/D or H" and "C or D/G or H").

9.3. As a final stage, the classes of non verbal explanation $D^0(R_i)$ (i.e. the modelling tasks) are specified as a Task Structure $TS(R_i)$. Canonically, these modelling tasks are represented as a command graph, which permits their concurrent execution. Since STATLAB is artificially constrained to accommodate only serial modelling operations, the program of Chapter 4 and Appendix E (which monitors the use of STATLAB) is, however, just as versatile.

9.4. Recalling the argument of Section 3.4, the fact that all of the primitive node relations are realised as properties of a modelling facility does not imply that the L^0 properties are comparable except insofar as they are related indirectly by entailment.

To emphasise the point suppose that node l and node p are primitive. It follows that R_l and R_p can be regarded as properties $\underline{\text{Prop}}$ l and $\underline{\text{Prop}}$ p (by anyone able to learn in R) and that there are $\overline{L^0}$ predicates $\underline{\text{Pred}^0}$l and $\underline{\text{Pred}^0}$p with values $\underline{\text{Val}}$ l 1 . . . . $\underline{\text{Val}}$ lm_l and Val p1 . . . . Val pm_p. The property values are certainly disjoint subsets and $\underline{\text{Prop}}$ l or $\underline{\text{Prop}}$ p may thus be regarded as sets (rather than 1-adic-relations) namely

$$S_l = \underline{\text{Prop}}\ l = (\underline{\text{Val}}\ l, 1)\ \cup \ldots \cup\ (\underline{\text{Val}}\ l, m_l)$$

$$S_p = \underline{\text{Prop}}\ p = (\underline{\text{Val}}\ p, 1)\ \cup \ldots \cup\ (\underline{\text{Val}}\ p, m_p)$$

But, for the reasons already rehearsed, it is neither necessarily, or usually the case, that S_l and S_p are, in any than an artificial sense, subsets of the same set.

9.5. As a very practical consequence, the objects picked out by L^0 statements in $\underline{\text{Pred}^0}$l (s) and $\underline{\text{Pred}^0}$p (s) (where s is an L^0 object variable) may belong to different parts of a modelling facility which (apart from their physical proximity) may as well be considered as distinct modelling facilities. For example, the left hand side and the right hand side of the STATLAB facility are distinct, in this manner. There is a very clear sense in which temporally or spatially demarcated occurrences (left hand side experimental results) are comparable with the abstract and temporal entities (the right hand side events) if and only if they are compared by the major (real, abstract) analogy of the theory embodied in the conversational domain. As a matter of fact the comparison is restricted, in any case, to ergodic or statistically stationary ensembles and it is contingent upon the profound but still tenuous notion of statistical dependency and statistical independence.

9.6. Moreover, it does not follow that a student A (who is able to learn in R) can only regard the primitive topic relations as properties. For example, if R_j is a relation between R_l, R_p (or $\underline{\text{Prop}}$ l and $\underline{\text{Prop}}$ p), A may also be able to regard R_j as a property. He can do so, on a given occasion n, if there is some $\underline{\text{Proc}}^0_A$j in π^0_A (n).

For example, the novice, learning probability theory, regards only the primitives as properties and consequently assimilates such concepts as "composite event" or, at a higher level in $D^1(R)$ "exclusive/inclusive/composite events" as relations between the initial properties. But any moderately versed student may regard "composite event" or "subset" as a property of the modelling facility and in the limit any student will be able to regard "probabilistic" as a property of general "inference" i.e. to understand probabilistic inference.

10. The L^1 description of R

Once that R is established, the IA labels its nodes using an index i = 1, ... i max and calls upon the source to give an L^1 description of R in which each node is distinct. This description is to be made in terms of unary but many valued L^1 predicates, called descriptors, and is a matrix that accommodates $D^1(R)$ (the descriptors are part of $D^1(R)$ but not all of it). It is assumed that any student can understand and determine the value of any L^1 descriptor. Hence, L^1 descriptors are primitives. But, as a convention to be adopted throughout this book (but rescinded in the next volume) they do not appear as nodes in R.

10.1. *Conditions to be satisfied by descriptors.* A descriptor names an L^1 predicate, $\underline{\text{Pred}}^1_\xi$ (e) where e is an L^1 object variable ranging over a set of place holders or storage positions in the interface at which $D^1(R)$ will be inscribed as ES(R) and $\xi = 1, \ldots c$.

As a first requirement, the source must furnish a *fine structure family* (Banerji 1970) of descriptors which are properties with values $\underline{\text{Val}}\ \xi\ 1 \ldots \underline{\text{Val}}\ \xi\ m_\xi$, each $\underline{\text{Val}}$ being a disjoint subset of place holders in the interface. If the descriptors are a fine structure family, then unique objects, in this case place holders, are designated by conjunctive expressions in all c predicates named by the descriptors. That is, if a and b are particular values, the L^1 expression $\underline{\text{Min Sent}}^1$ (e) $\equiv$ ($\underline{\text{Pred}}^1_1$ (e), = $\underline{\text{Val}}$ 1, a) $\wedge \ldots \wedge$ ($\underline{\text{Pred}}^1_c$(e) = $\underline{\text{Val}}$ c, b)) is a unique *place holder.* If the fine structure family is full with respect to R then the place holders are nodes, labelled i by the IA; as a result of which the object variable e can be replaced by the index i. However, the source is required to specify a fine structure family that is not full so that some $\underline{\text{Min Sents}}$ describe place holders that are not currently occupied by nodes, but which may become occupied. Further, so that certain of the *values* of different $\underline{\text{Pred}}^1_\xi$ may be comparable, the object variable is taken to range over a product set "place holders × integers" rather than a simple set "place holders".

10.2 *Head nodes and induced descriptors.* In order to test an entailment mesh for cyclicity, a head node must be chosen. This choice necessarily induces a descriptor, call it the primary descriptor, and name it $\underline{\text{Pred}}^1_1$, conveying the general meaning "degree of superordination/subordination" in terms of depth (arc distance) d, of a node from the head node.

The induced descriptor names an integer valued predicate defined on place holders × integer. Unless the pruned entailment mesh is a chain, values of this predicate alone furnish only a partial description of nodes in the mesh; for example, construct any tree and notice that ambiguity occurs at its bifurcation, hence, other descriptors are needed.

From the definition of cyclicity, if the entailment mesh is cyclic, then at least one node in the pruned entailment mesh has more than one incident arcs from other than distinguished nodes. Hence, this node is the root of a tree (a perverse but usual notation). Hence, if the original mesh is cyclic, then the pruned mesh contains at least one tree (so it is not a chain) and the primary descriptor yields an incomplete description.

It follows that the requirement of complete description (being able to ostend each node uniquely) calls for descriptors over and above the primary descriptor.

In the main example, (the CASTE interface and the conversational domain of probability theory) the primary descriptor is 2nd and the set of descriptors 1st, 2nd, 3rd, 4th is a fine structure family.

10.3. *Redundant descriptors.* Although a source is required to select at least one fine structure family which is not full and although (Section 10.2) this means choosing more than one descriptor, he may, if he wishes, select more than one fine structure family, and/or descriptors that do not belong to a fine structure family; either expedient yields a redundant description.

10.4. *Eliciting descriptors.* The act of choosing a head node induces the primary (superordinate/subordinate) descriptor. Other descriptors may be invented spontaneously, but if not, they are elicited by one of the following techniques.

Suppose that a given value of $\underline{\text{Pred}}^1_1$ (e) (the primary descriptor) is unique (for example, on a single head node). If so, no further descriptors are required for this value (though there is no objection to adding any that are proposed).

Suppose there are just two nodes, i, and j, that have the same value on $\underline{\text{Pred}}^1_1$ (e). A further descriptor is needed and is obtained by encouraging the source to discriminate node i from node j according to some property which is named as a unary L^1 predicate $\underline{\text{Pred}}^1_2$ (e) with distinct values on node i and on node j.

Suppose there is at least a triad of nodes i, j, k, that have the same value on $\underline{\text{Pred}}^1_1$ (e) then the ranking of $\underline{\text{Pred}}^1_2$ values is continued over all such nodes i, j, k, l, m, etc. and the appropriate values of $\underline{\text{Pred}}^1_2$ (e) are assigned to these nodes.

This process is repeated for all values of $\underline{\text{Pred}}^1_1$ (e) to generate further predicates $\underline{\text{Pred}}^1_\xi$(e). If there is a previously elicited predicate $\underline{\text{Pred}}^1_{\xi-1}$ (e) then its property is rated with respect to the nodes picked out by each new value of $\underline{\mathrm{Pred}}^1_i(e)$, thus creating further values of $\underline{\mathrm{Pred}}^1_{i-1}(e)$.

By iterating these procedures it is possible to assign real descriptor values to all of the nodes existing in the pruned cyclic entailment mesh and consequently to build up a fine structure family of L^1 predicates. In general, the fine structure family is not full. If it is, the descriptors must be augmented so that there are objects (place holders, storage positions) pointed at by Min Sent $\underline{L}^1$s, that are not existing nodes (indexed by i). These are, however, nodes that the source is able to appreciate, insofar as he can describe them in relation to the nodes in the currently specified mesh. Moreover, he may (or may not) be able to cite the topic relations that correspond to these L^1 object descriptions. But if he can do so he is able to extend the original relational network.

10.5. *Status of descriptions.* L^1 sentences in the $\underline{\mathrm{Pred}}^1(e)$ (including Min Sent's as a special case) refer to classes of knowables. Further, expressions in many place L^1 predicates signify relations between knowables. They do not, directly, refer to the topic relations themselves i.e. L^1 statements are about the "cognitive map" in terms of which either the source or any student who can give a meaning to the descriptors (in the sense that values of $\underline{\mathrm{Pred}}^1_i(e)$ mean subordinate/superordinate) is able to appreciate a topic without necessarily, at this stage, knowing what that topic is.

The IA insists that each descriptor is given a meaning and encourages the source to use meanings that are familiar to the students he has in mind.

10.6 *Adding the operator data base.* Each node in the IA's L^* description of R is associated with an ordered set of relational operators. These are represented in L^1 as an operator data base (formally, in terms of many place L^1 predicates and a primitive "apply" which is associated with each pair of Min Sent's).

10.7. *Forming $D^1(R)$.* To each node (which may be identified in L^* by an index i and in L^1 by a $\underline{\text{Min Sent}}^1$) the IA assigns a set of variables, the L^1 marker predicates, which signal the state of a node during a strict conversation. The L^1 description $D^1(R)$ which is inscribed at the interface as ES(R) consists in the L^1 descriptors, the L^1 translation of the operator data base and (for each node), the marker predicates.

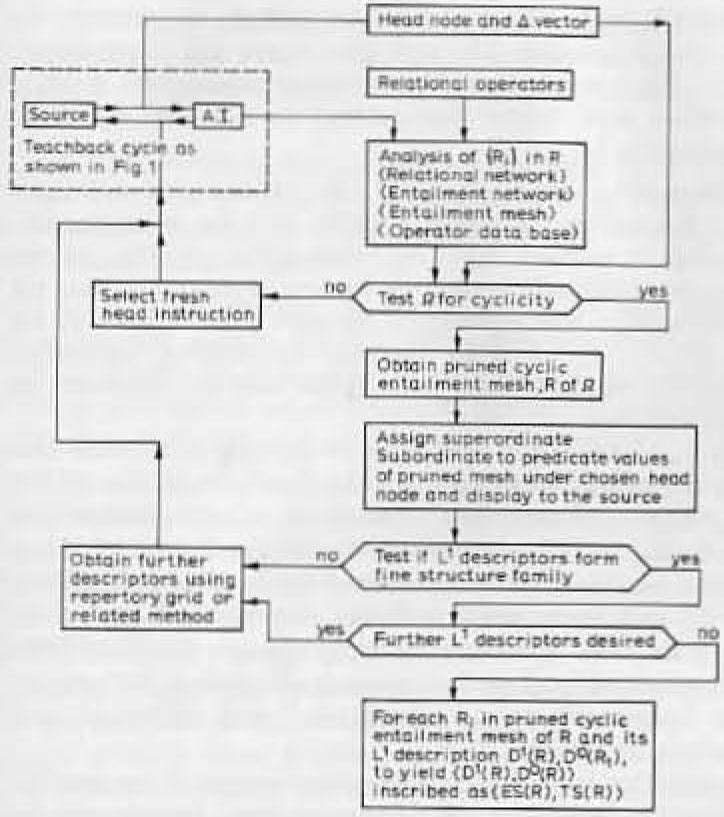


Fig. 7.22. "Flow Chart" or "Process Chart" approximating the course assembly and pruning heuristic.

10.8. *Summary.* At this point, it is possible to interpret all terms in the flow chart of Fig. 22, which summarises the foregoing discussion of how to construct a conversational domain.

11. Overview of some extensions of Entailment Meshes and the Course Assembly Process.

There is no uniquely "best" description of the entailment mesh (Ω) that develops during the source/IA dialogue. The source periodically announces a head (perhaps multiple) under which Ω is tested for cyclicity and, if cyclic, is pruned. Choice of a head generates the primary L^1 descriptor ("superordinate" or "subordinate") but others must be added to satisfy the requirements of Section 10, and the source has considerable freedom in this respect. Moreover, he could always have selected several distinct head topics (thus setting R, in $D^1(R)$ or ES(R), equal to a specific Ω_o.

11.1 *Alternative descriptions of an entailement mesh and redundant descriptors.* Although CASTE, as it has been detailed, only presents a student with one description, $D^1(R)$, of the pruned cyclic entailment mesh, the system actually contains the data represented in the entailment network. Hence, even in the existing system, it is quite possible to employ redundant descriptors on the same mesh with values that are uncovered by explore transactions.

In addition, there are several ways of extending the collection of descriptors and increasing the possible level of redundancy. For example, choice of additional L^1 predicates or of a different fine structure family of L^1 predicates both achieve this result. More fundamental modifications are produced by choosing further head nodes in the entailment mesh, with the property that the pruned mesh is cyclic. This operation certainly changes the description because it necessarily alters the meaning attached to the primary descriptor "superordinate and subordinate" with respect of some (and usually many) nodes.

This possibility is especially interesting because it appears that one outstanding parameter of a conversational domain and the underlying collection of knowables is the number of head topics which can be chosen with the property that all of them yield a cyclic and prunable entailment mesh (since each choice gives rise to a different primary descriptor on a given mesh Ω).

11.2. *The "Quality" index.* Let HedNum (M) stand for the number of other-than-primitive or distinguished-nodes in Ω and call the quality of this mesh Qual (Ω), namely

$$\underline{\text{Qual}}\,(\Omega) = \frac{\underline{\text{HedNum}}\,(\Omega)}{\text{Number of head nodes yielding a cyclic mesh, R in } \Omega}$$

Thus, Qual (Ω) is the ratio of nodes that could, in principle, count as head nodes to the number that give rise to meaningful structures.

11.2.1. Qual (Ω) can equal 1 (maximum quality) as is shown by constructing any closed definition of a relational operator, (say, of natural composition in terms of other operators here, natural join and projection).

11.2.2. Obviously Qual (Ω) = 0 if HedNum (Ω) = 0 because Ω does not yield a cyclic structure after pruning. There are certain entailment meshes for which HedNum (Ω) = number of (non primitive) nodes in mesh.

11.2.3. The motive in defining Qual and HedNum is conceptually simple, but has widespread ramifications. HedNum (Ω) indicates the number of topic relations in Ω that could be named and that, if named, could be reproduced as memories: Qual (Ω) the proportion of nodes that have this property.

11.2.4. The larger the value of Qual (Ω), the more richly are the topics "associated". But notice that "association" (as normally used in psychology) fails to indicate the information preserved in linking R_i and R_j. Cyclic entailment does so and Qual (Ω) is thus an index of how much potential information is dispersed in the mesh.

11.2.5. Qual (Ω) may also be regarded as a measure of the L-explicability of the set of topic relations in Ω; the number of ways in which these relations can be explained. It may be possible to replace this index by a logical content or conversely a logical information measure. As an optimising principle, for the source/IA dialogue, we recommend building relational networks, meshes, etc. to maximise the value of Qual (Ω).

11.2.6. For any mesh Ω it is possible to generate HedNum (Ω) distinct descriptions of the conversational domain by choosing this number of different head node(s) and deriving the pruned cyclic meshes), R.

If the operations discussed in Sections 8 and 9 are carried out, then each R is given at least one description $D^1(R)$ and it may be given several, (in principle, it can be given an indefinitely large number of them since the source is free to select any fine structure family of predicates $P^1_\xi(e)$ and to add redundant descriptors if he wishes). These descriptions are indexed θ (as the $D^1_\theta(R)$ and $D^1(R)$ (unindexed) may either be used to refer to one of them or the entire class. By convention, just K (R) descriptions are, in fact, specified by the source (hence $\theta = 1 \ldots . K(R)$). The K (R) descriptions D^1_θ form a family the "fundamental description class of R" in which the primary descriptor (for superordinate/subordinate) has the same meaning.

11.3. *Some properties of alternative descriptions.* The CASTE facility is undergoing fairly rapid development in order to accommodate the features noted in the last few paragraphs.

Clearly, it is possible to store any finite number of fundamental descriptions $D^1_\theta(R)$ of R and the future versions of CASTE will be provided with means for accessing them.

It is also possible to store, for any Ω, the HedNum (Ω) pruned cyclic meshes R and to access the corresponding $D^1_\theta(R)$. Moreover, there is no reason why further descriptors should not be added to the list (the selection of K is arbitrary) in which case $K = K_n(R)$ and $\theta = 1, \ldots. K_n(R)$ (n, as before, an occasion index).

Since entirely different additions and modification are carried out by an evolutionary course assembly heuristic (when $R = R(n)$) it is important to recognise the limitations involved. Choice of fresh descriptors (increasing $K_n(R)$) does not alter HedNum (R): in general, evolution does so. Nor does a change of description, $D^1_\theta(R)$ into $D^1_{\theta+1}(R)$ alter the head in R; still less, the structure of R. In general, evolution does both. Clearly, there is a connection between describing R afresh and the act of building R (Section 8 and Fig. 22). Description is a prerequisite for meaningful construction. But the two are not identical and describing a topic relation differently does not, on its own, change the relation in question.

11.4. *Student accessing facilities.* As CASTE has been described the student is presented with one $D^1_\theta(R)$. He can, however, gain access to further descriptions in the same fundamental class by explore transactions. Though the number of descriptor values that may be explored (ultimately to generate all of $D^1(R)$) is restricted by the current equipment the limitation is in no way essential.

In the next version of CASTE a student will be provided with two facilities. On the one hand, he can display any of K(R) distinct $D^1_\theta(R)$ as a whole and may also display any pair of them $D^1_{\theta_1}(R)$, $D^1_{\theta_2}(R)$ for comparative purposes. Since these operations change neither the head node nor the set of nodes in R, the marking (with L^1 predicates such as goal or aim or understand) is unchanged. On the other hand, the next version of CASTE also allows the student to access the HedNum distinct R (if more than one exists) and, for each R, to gain access to K(R) distinct $D^1_\theta(R)$. Further, in respect of each R, he may opt to learn; and, if he does so, the marker states are changed.

Some technical complications arise, because of that. For each head (each R) it is necessary to store all of the marker distributions produced by learning under that head (some and only some of the marker values are common to several heads). The problem is by no means insuperable (HedNum is modestly sized in the cases so far encountered). The additions to the facility have value both as a method for achieving genuine multiple aim operations and as a practical aid in learning.

11.5. *Operating on the descriptors and their values.* A student using this new equipment, will be able to specify descriptors of his own (and thus add to the data available) provided he can give values to these descriptors. He may be encouraged to select descriptors by questions like "do you see the endomentrium as one regulator in the menstrual cycle" or "do you see the menstrual cycle as a process required to keep the endometrium in working order?". He may also be assisted in selecting descriptors by an easily mechanised repertory grid facility (nodes are the objects; the constructs elicited, and evaluated on each node, are the fresh descriptors).

The crucial point is that none of these operations necessarily alter the student's role; A in other words, may still converse as a student in the same domain and about the same subject matter. His fresh ways of describing the topic relations (as values of freshly specified L^1 descriptions) are not allowed, under these circumstances, to change the modelling facility, the task structure, or any $D^0(R_i)$. That is, the student is not allowed to change the knowable subject matter or to explain the R_i differently though he may come to appreciate the subject matter in question from an entirely different perspective.

11.6. *Relaxation of constraints.* The "not allowed" clause in the last subsection is an arbitrary directive, imposed for convenience, only. If all of the $D^1(R)$ for each of the HedNum R are stored at the interface (as they are in the currently designed and soon-to-be operating system) then, at any occasion n in a strict conversation across the interface, there is no mechanical reason why a student should not select another head node.

If a subject matter expert chooses a new head in Ω (which is checked as cyclic and prunable) he would not change the number of other than initial nodes but he generally would change the meaning of a given topic relation R_i; that is he would be likely to specify a new $D^0(R_i)$ or to request new explanations of R_i.

These comments apply with just as much cogency if the HedNum possible heads in Ω are generated and the subject matter expert selects amongst them.

The possibilities are restricted but not otherwise altered by the requirement that any new explanation demanded can also be constructed (as a model) in the modelling facility already provided. This clause is fairly restrictive if the facility is STATLAB (though, within the meaning of the clause, it is possible to model several subjects even on STATLAB; for example "probabilistic automata" rather than "statistical experiments"). The clause is far less obtrusive provided the modelling facility is liberal enough (for example, if it is an easy-to-use programming language and models are programs).

11.7. *The Student acting as a subject matter expert.* There is nothing to prevent a student doing any or all of these constructive operations, usually the prerogative of a source or subject matter expert, if we relax the arbitrary constraint that "students must not fiddle with the subject matter".

If we do relax that constraint, if the node index is "open ended" ($i(\max) = i \max(n)$), and if the heuristic of Fig. 22 is applied to any Ω for any putative head node; then there is nothing to prevent a student doing the other trick of a subject matter expert; namely, adding nodes for legal (i.e. cyclic and consistent) topic relations, so that $R = R(n)$.

Under these circumstances it is tidier to say the student "changes role" and "becomes a temporary subject matter expert" or, vice versa, the subject matter expert becomes a learner (as the source does, in teachback interactions with the IA).

The crucial point is that when $i\,(\max) = i\,\mathrm{Max}(n)$, the heuristic of Fig. 22 can still be executed in the CASTE system as it stands, as described in the next chapter where the operating routine is listed. Hence a student can build conversational domains, or vice versa, the source can replace a student in the system.

11.8. *Technical modifications.* The existing procedures (for checking cyclicity, for pruning, and so on) are quite adequate. But the process of course assembly is rendered cumbersome by purely

technical limitations like the fact that topic names are written manually and that new nodes are spare locations on the display fascia which have to be labelled as the topic relation they represent.

Most of these problems are surmounted and hardly worth discussion. The technical innovations needed to instrument the facilities of Section 11.6 and Section 11.7 include arrangements for displaying nodes and their connections on a battery of electronic display tubes; for reading existing entailment structures from store onto the display tubes and for operating upon them. The task of making finite additions is quite trivial provided the additions are legal, which is guaranteed by the existing tests for the consistency of the mesh and cyclicity of the mesh.

12. Other Methods of Representing Knowables

Structures that superficially resemble pruned entailment structures have been exhibited from time to time, ever since Gagné's original work on learning to complete and find formulae for number series. Familiarity, perhaps, breeds a kind of contempt; at least it is common to find quite indefensible "knowledge structures" that may represent almost anything; an author's idea of how to teach, or learn, or arrange or display or describe whatever the linked nodules stand for (and this varies with the mood of the moment; sometimes they stand for concepts, sometimes classes, sometimes inscriptions and sometimes the section headings of a book). Such "structures" are often misleading when taken at face value.

12.1 *Some non-trivial schemes.* Some "knowledge structures" are perfectly genuine and have proved to be informative and useful. All the acceptable structures appear to be encompassed as special cases within the theory of entailment meshes and/or the partial theory of task structures. But it is illuminating as well as proper to review other constructions and the different meanings that are assigned to the nodes, links, etc. of which they are compounded.

12.1.1. Gagné's work in this field has historical priority and is intrinsically interesting. For the number series task (alluded to earlier) Gagné (1962) produced an hierarchy of nodes under the superordinate/subordinate relation called an hierarchy of knowledge. Each node stands for a task class; it may also, on slight reinterpretation, stand for a topic relation. The operationally crucial issue is that if A is superordinate to B and C then learning A depends, as a condition, upon knowing both B and C. Using appropriate randomised tests for comprehension, Gagné found that the following predictions were confirmed in a virtually all or none fashion.

Inference Number	Student with knowledge (+) or ignorance (−) of a topic listed below.	Will give a positive (+) or negative (−) or an undetermined (o) test result in respect to the related sub-tasks listed below.
1	Superordinate (+)	*All* Subordinates (+)
2	Superordinate (−)	Subordinates undetermined (o)
3	*All* subordinates (+)	Superordinate (+)
4	*Any* subordinates (−)	Superordinate (−)

Especially in his later work (for example, in Nice, 1968) Gagné takes an electic point of view, without commitment to what knowledge really is. However, at least for the number series project, it is possible to interpret the nodes as relations such that any superordinate relation contains the explanation of all its subordinates. Hence, the entailment mesh of these relations is cyclic and, if pruned, one of its kernels is the exhibited knowledge structure. Students were restricted to this kernel (at least for their training in later parts of the experiment). The very definite results obtained may be due to the fact that this structure *is* a valid entailment mesh; these inferences can be made, with confidence in their empirical confirmation, for any fully conjunctive part of an entailment mesh, such as a single kernel.

Without cyclicity, there is no reason inherent in the mesh, why concepts for all subordinate topics should be reconstructible given knowledge of a superordinate (Inference 1) as, in fact, they are.

Inferences 3 and 4 are straightforward, since an ability to discover relations is assumed, quite explicitly, in Gagné's framework. In our own system, it is necessary to introduce a distinguished node at each confluence of the mesh and to assume that L^1 operators $\underline{\mathrm{Proc}}^1_A$ exist in π^1_A for all of the constructions required.

12.1.2. The difficulties over a broader interpretation of Gagné's system are as follows: (a) It exhibits (unequivocably) only single kernels of a head node even if nodes are assumed to represent reproducible relations. For the example given this is probably quite unimportant since most students (by dint of common schooling or the like) did explain number series formulae from the basis exhibited. In general, of course, there are many ways of getting to know and the permissible inferences about subordinates given knowledge of a superordinate are disjunctive in form. (b) As a result of suppressing disjunction, analogies between aspects of the subject matter can only be represented in an ad hoc fashion. (c) It is not clear whether the structure is a restricted entailment structure (like ES(R)) or whether it is a restricted task structure TS(R) (or $TS(R_i)$ for each R_i in R).

12.1.3. Due to the language and nomenclature of "behavioural objectives", which is the idiom of the paper, this distinction is glossed. But it is brought into the foreground by considering different potential applications of the knowledge structure; namely, as a prescription for how the knowledge is knowable; (ES(R) definitely); as a prescription for how to bring about the relations in question (TS(R) definitely); or as a prescription for how best to teach.

It is particularly important to stress the last of these applications because it will turn out that learning/teaching strategies cannot generally be represented in a strictly hierarchical system, though some strategies can be.

12.1.4. These areas of application and concern winkle out the foci of other investigator's attention. For example Cox et al. (1972) are truly concerned with structures like ES(R) in which the relational operators are the mathematical operators, integration and differentiation and the nodes represent classes of identified mathematical expressions (for force, velocity momentum, acceleration). Their structure, for one aspect of classical physics, is sketched in Fig. 23 (since the original paper is mineographed). The

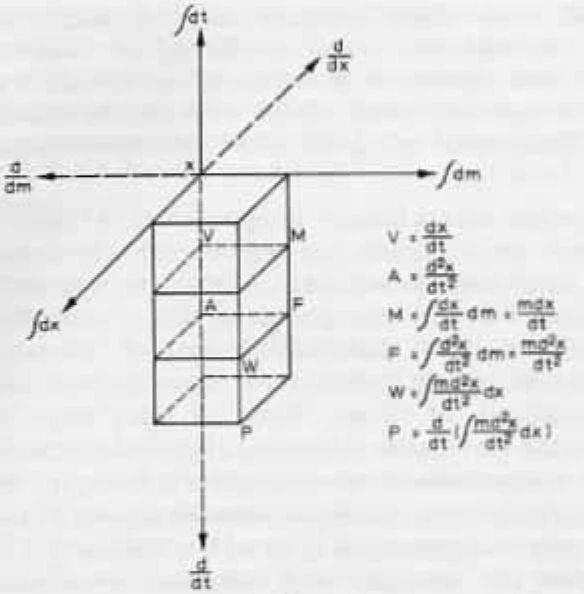


Fig. 7.23. Geometric model of selected concepts in analytical mechanics.

continuities required to convert a relational network into a space exist in this particular subject matter area. But this circumstance is rare and there is one aspect of Fig. 23 that is a little deceptive; all analogies that identify the relations (expressions) are omitted and assumed to exist, a trick comparable to pruning (Section 8). The difficulty engendered by suppressing them becomes evident in view of the author's later finding that a structure works *best* as a tutorial prescription when the topics are "richly connected" (like physical ideas) rather than "left uninterpreted" like categories of abstract designs. Fig. 24 (our own responsibility only) is a superimposition on Fig. 23 of the analogy relations most obtrusive in Cox et al.'s graph, which are needed to furnish the cognitive adhesion they find in their experiments and we confirm, as a dominant result, in our own.

12.1.5. Bunderson (1967) has experimented with various kinds of structure and applied them both widely and successfully. In practice (and a deeper meaning exists) Bunderson's subject matter structures are either task structures or else statements of how the course modules fit together. Some possible difficulties over the

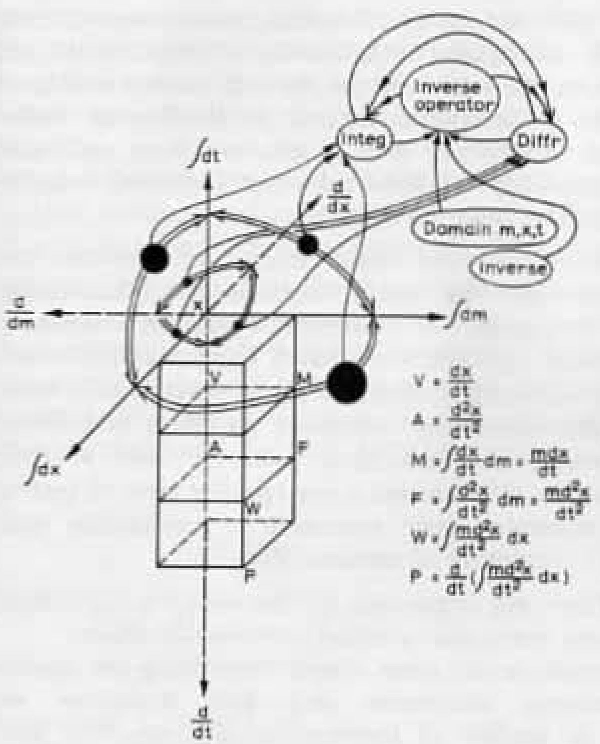


Fig. 7.24. The structure in Fig. 23 together with certain of the analogy relations that are implicitly stated.

former interpretation are rehearsed in Chapter 9. Since the latter interpretation amounts to a condensed statement of a careful and rather elegant tutorial design process, these structures have definite utility and are not at all open to the criticism levelled in the preface to this section.

12.1.6. Horn (1971) has developed a very elaborate system called "information mapping" which seems to have the status of one description (or possibly several descriptions) of a tacitly assumed entailment mesh. This remark may understate the real calibre of information maps. The work, however, came to my attention only recently and, after a couple of discussions with Horn himself, this appears to be a fair summary of the position. If so, the main weakness of information maps (which can certainly be remedied) is that there is no explicit theory for generating the underlying mesh. One serious and obvious consequence is that canons of what is memorable or knowable or learnable do not emerge directly from the structure (as they do via the relational operators and the P Individual model, in the theory under discussion). If such criteria are used at all, then they rest upon assumptions or empirically validated rules that are added onto the framework in which the "map" is made.

12.1.7. Apart from the work under discussion there are two comprehensive attempts to unify notions of "knowledge structure" with categories of "mental operation" whereby concepts may be built up. On the one hand, Landa, (1971) who is working in Russia, conceives a structure as the ossature of a body of instructional algorithms and heuristics; Kopstein and Seidal (1965–1971) working in Washington take (at least as their working hypothesis in this matter) a less specific view of mental operations. For example, their approach is compatible with Guildford's (1956) "Structure of intellect Model".

12.1.8. Both efforts are impressive, in line with the application aspect of the present work, and oriented towards education.

Relational networks much more closely resembling the meshes underlying entailment structures and task structures are commonly used in studies of machine intelligence but, until recently, have seldom been given a direct psychological or educational interpretation. Two outstanding exceptions are provided by the work of Norman and his group (1973) and by the independent work of Scandura (1973).

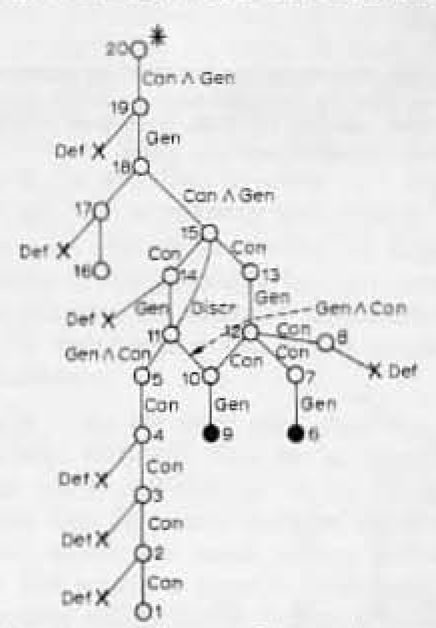
These two exceptions are mentioned, but not discussed, because I only learned of them in detail as this book was going to press. Norman and his colleagues have a comparable view of memory as the reconstruction or reproduction of relations. A similar theme is developed (at any rate in personal discussion) by Scandura. He also stresses the open ended or evolutionary character of a competent representation and his book, Scandura (1973), in many ways parallels the theory of knowables presented in this chapter, though the notation and mechanisms of his theory are rather different.

12.2. *The deceptiveness of simplicity.* Overall, there is a prevailing and damaging confusion between noun hierarchies, (ordered under class inclusion or the superordinate/subordinate dimension) and verb networks (relational networks from which,

for example, it is possible to derive entailment structures). This confusion leads to a number of half truths. For example, "It is easy to structure or subject matter (how? what is a subject matter? why?) provided that the subject is well specified (how?)". As a corollary to this assertion, the fairly sophisticated construction process described in this chapter is superfluous; moreover, if the subject matter is, by some criterion or other "well specified" there is nothing remarkable about being able to structure it.

12.2.1. It is easy to draw structures (the simplest of them are trees) that represent collections of nouns i.e. topic names for any well defined subject matter. Conversely, it is hard or impossible to draw noun structures for badly or hazily defined materials.

Thus, Fig. 24 shows a taxonomic structure used by Salazar, Resines and myself (1965, 1971) in order to design programmed



Type names	Other symbols
Def : Definition	○ : Subconcept node
Con : Concatenation	● : Terminal node denoting initial subconcept
Gen : Generalisation	✱ : Concept at head of ordering
Discr: Discrimination	X : Terminal branch for introducing definition

Fig. 7.25 (0). A concept ordering for a logic programme: The numbers are assigned by a heuristic to indicate sequence of instruction.

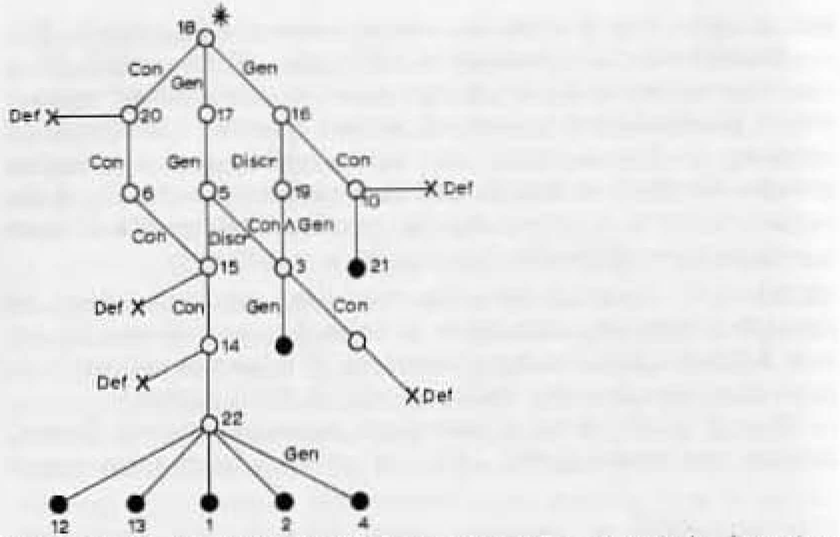


Fig. 7.25 (i). Alternative knowledge structures for concepts in elementary logic: Numbers 1—20 are thore assigned by heuristic used to define tutorial order in Fig. 24. No order has been assigned to these structures. Numbers above 20 designate nodes that do not appear in previous figure.

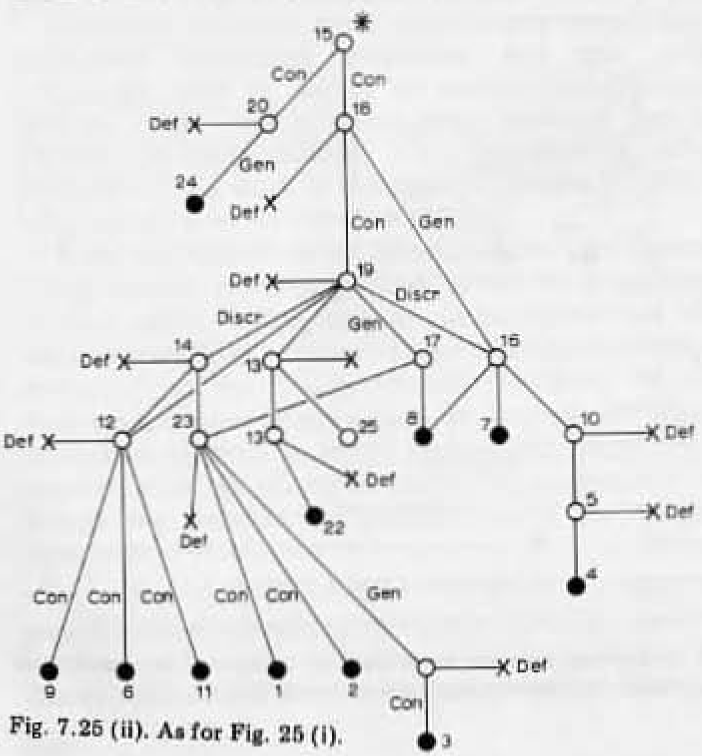


Fig. 7.25 (ii). As for Fig. 25 (i).

text materials for symbolic logic. When the critics say, "It is easy to draw such a structure because logic is well specified" we concur; but point out that indefinitely many structures are possible. We used several dozen for logic. Fig. 25 shows a few of them.

12.2.2. The trouble begins as soon as integrating operations are introduced (in Fig. 25 (o), concatenation, generalisation and abstraction) which serve as names for loosely conceived modes of discovery. The operations are really tagged onto the structure, after it is built. If nouns a and b are given (as the topic headings of Fig. 25, for instance) these "operations" (more properly "operation tags") are adjoined in an attempt to describe how a student gets from a concept of a to a concept of b. But, notice, a and b are structured first.

True, the operation tags are not valueless; for example, guidance can be furnished in terms of them. But the criteria for assessing the adequacy of a structure as representing a knowable (that is, something that is a memory or even, harking back to outdated theories, a "trace", that does not interfere with its neighbours) cannot be specified in terms of a taxonomy and its labels. For example, the consistency and cyclicity requirements of the present discussion cannot be erected in these terms. At the most, it is possible to use plausible but not very well supported heuristics, (for example, the order of presentation heuristic, noted in Fig. 25 (o)) to determine points where students are likely to encounter difficulties.

12.2.3. As we have seen, building an entailment structure is relatively difficult. But, so far as our experience goes, it is not much more difficult to do so for a sloppily specified subject matter than for a well specified subject matter. In fact, if a closed, cyclic, entailment mesh exists, then something definite is being said. The source is literally forced (by the teachback dialogue) to propound an explicable theory he believes in or a set of alternative theories that are contrasted within a wider theory. There may be subject matters impervious to rational penetration (though I doubt it). If so they are uncommon for the following reasons.

(a) The source's definite theoretical statements need not be true (or even generally accepted); they are his coherent innovations.

(b) Theories may be related to one another by theories of a different type; for example, two scientific theories can either be compared within a wider scientific theory, or compared on historical grounds.

12.3. There is an obvious sense in which the teachback of source/IA dialogue impels the source to output a certain type of product. This constraint can be interpreted as restricting what the source really wants to say or compelling the source to commit himself in respect of explicable and teachable entities.

12.4. There is also a (somewhat less obvious) sense in which the source is required to look ahead and take a broad enough perspective to accommodate the kind of explanation he believes to be needed (that satisfies him, that he can justify as a head topic). Succinctly, the source must anticipate and entertain a genuine theory and in order to do so he (like a student learning the theory) must appreciate a description of the topic relations before these topics are stated. Though clear enough on watching the source at work, these facts are obscured by the linear-looking heuristic.

12.5. It is quite possible to represent simplified theories, provided that their limitations are revealed. It is also possible to theorise about descriptions (the same thing, perhaps). But the chatty dilution of a theory does not yield a cyclic entailment mesh and could not be represented by an entailment structure.

12.6. So, in practice, the source usually finds it convenient to delve into a subject at some depth. For example, in explaining multiplication and exponentiation, a source (following one of Goodstein's essays) may opt to show addition, multiplication (iterated addition), exponentiation (or iterated multiplication) . . . as members of a series recursively explained by the function

$$O_{n+1}(a, b+1) = O_n(a, O_{n+1}(a, b))$$

For which the base is

$$O_n(a, 1) = a \text{ for } n > 1$$

$$O_1(a, 1) = a + 1$$

$$O_0(a, b) = b + 1$$

To do so, the source must pave the way by defining "recursion" and "mathematical induction" (the special case for inherited properties;" if n has p and n + 1 has inherited p, then all successors have p"). He must also define "+" and "x" but he may as well

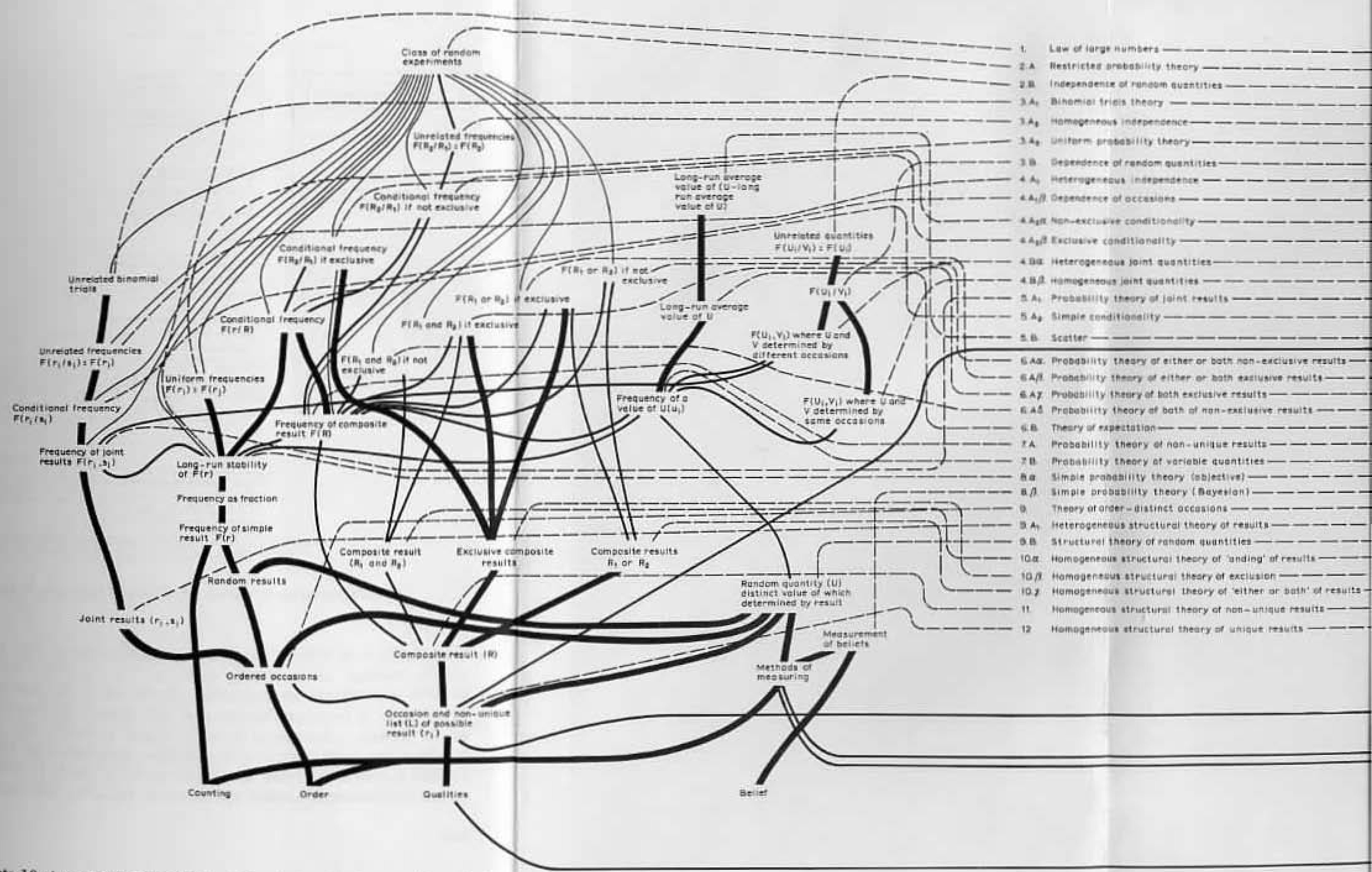
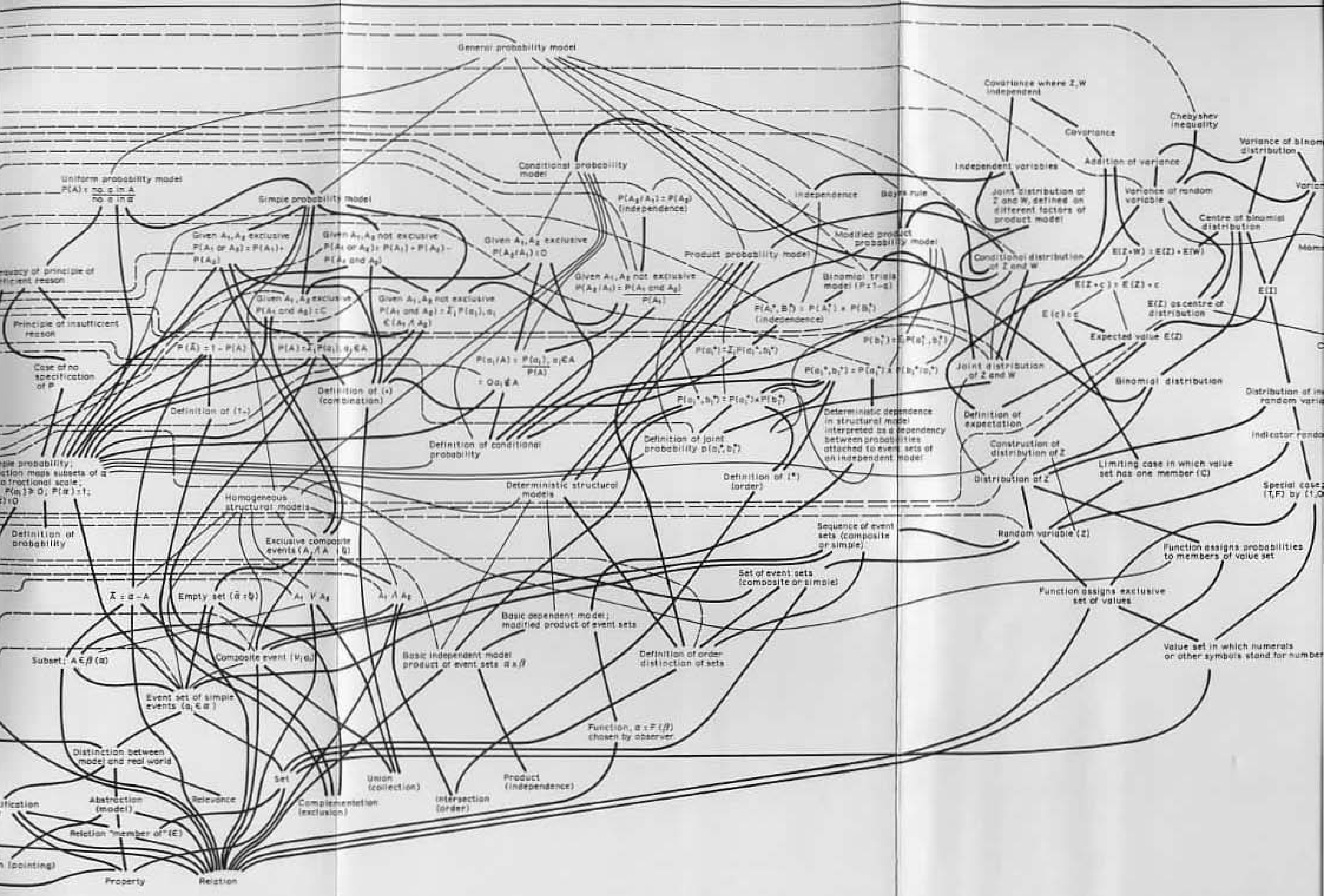
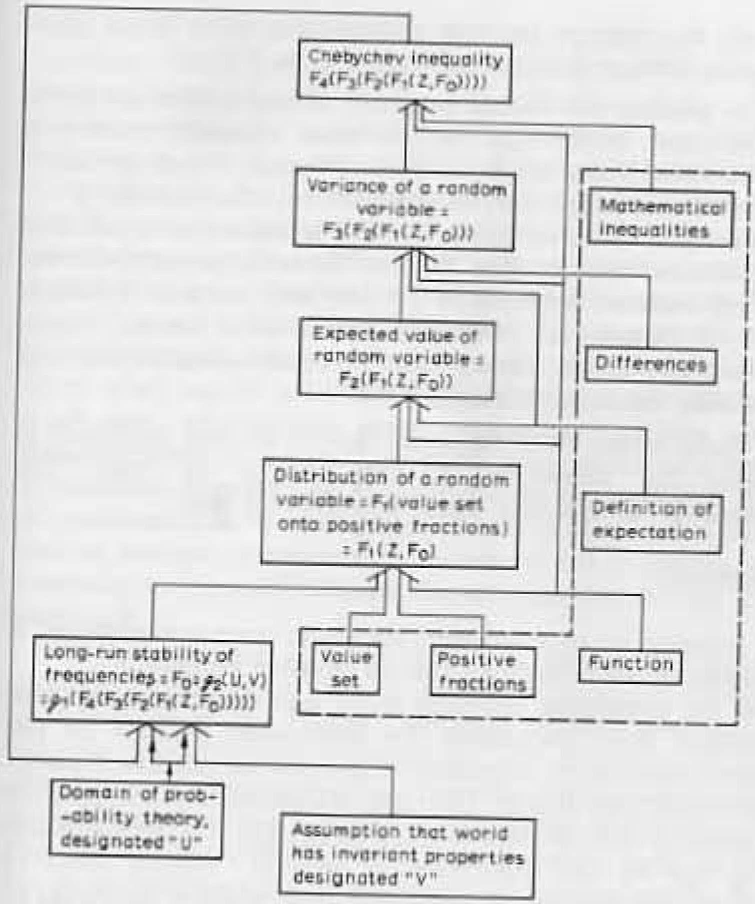


Plate 10. A pruned and described entailment structure (representing a thesis on part of statistics) in which the structure in Fig. 2 is embedded.



General probability model
Covariance where Z,W independent
Covariance
Chebyshev inequality
Addition of variance
Independent variables
Joint distribution of Z and W, defined on different factors of product model
Variance of random variable
Centre of binomial distribution
Variance of binomial distribution
Conditional probability model
Uniform probability model
P(A) = no. α in A / no. α in Ā
Simple probability model
P(A₂/A₁) = P(A₂) (independence)
Independence
Bayes rule
Modified product probability model
Product probability model
Conditional distribution of Z and W
E(Z+W) = E(Z) + E(W)
E(Z+c) = E(Z) + c
E(Z) as centre of distribution
E(c) = c
Expected value E(Z)
Given A₁,A₂ exclusive P(A₁ or A₂) = P(A₁) + P(A₂)
Given A₁,A₂ not exclusive P(A₁ or A₂) = P(A₁) + P(A₂) − P(A₁ and A₂)
Given A₁,A₂ exclusive P(A₂/A₁) = 0
Given A₁,A₂ not exclusive P(A₂/A₁) = P(A₁ and A₂) / P(A₁)
Binomial trials model (P = 1 − q)
Principle of insufficient reason
Given A₁,A₂ exclusive P(A₁ and A₂) = C
Case of no specification of P
Definition of (+) (combination)
Definition of (1−)
Definition of conditional probability
Definition of joint probability
Definition of (*) (order)
Joint distribution of Z and W
Binomial distribution
Definition of expectation
Construction of distribution of Z
Distribution of Z
Limiting case in which value set has one member (C)
Deterministic structural models
Deterministic dependence in structural model interpreted as a dependency between probabilities attached to event sets of an independent model
Homogeneous structural models
Definition of probability
Exclusive composite events
Sequence of event sets (composite or simple)
Random variable (Z)
Function assigns probabilities to members of value set
Set of event sets (composite or simple)
Function assigns exclusive set of values
Empty set
Basic dependent model; modified product of event sets
Basic independent model product of event sets
Definition of order distinction of sets
Value set in which numerals or other symbols stand for numbers
Composite event
Event set of simple events
Function, α = F(β) chosen by observer
Distinction between model and real world
Set
Union (collection)
Product (independence)
Abstraction (model)
Relevance
Complementation (exclusion)
Intersection (order)
Relation "member of" (∈)
Property
Relation



Enclosed nodes contribute Z which is isomorph to relation that constrains "U". $\not{g}_2$ is a join, $\not{g}_1$ is a restriction

Fig. 7.26. Explanation scheme in the domain of elementary statistics.

define "+" recursively using "successor". Once it is pointed out that "+" or "x" or any other iterate is an instance of the O-function, all of the specific operations are explained in terms of the O-function (though they may also be explained in other ways).

Fig. 26 shows a similar type of explanation scheme in the domain of elementary statistics; it is derived from the relational network of Plate 10 (where, however, only the pattern is visible). Amongst the primitive terms in this scheme is probability theory (in exactly the sense of the main example for which ES(R) (Plate 3) is derived from a cyclic subnetwork of Plate 10).

12.7. In general, the Gestalt property of a consistent and cyclic entailment mesh is two-fold. (a) The nodes represent knowables; either known or learnable. (b) If such a structure is learned then it may serve as a primitive node in any other entailment mesh.

For example, if "probability theory" is named as a primitive node in a mesh based on Fig. 26, then the entity concerned is the entire topic relation learned if (R), of the main example, is learned (plate 3 of Chapter 4). Mastering "probability theory" means "explaining probability theory" and "reconstructing the concept of probability theory in at least one way".

12.8. As a remark to be taken up later, a student may learn to reconstruct his knowledge in any number of ways. The requirement is "at least one", not "one and only one".

13. Questions

The potency of an entailment structure, once it is obtained is exemplified by the following notes on question generation.

Given the entailment structure and a task structure competent to delineate Base Commands for each topic relation in the entailment structure, it is possible to generate all possible PQuest^0 to cover the subject matter. The rules of construction are specified in Appendix D (they involve forming an Alt Set of alternatives and selecting one of the alternatives as correct). Various kinds of PQuest^0 are possible for R_i depending upon whether the values of one or more of the subordinate relations are initially defined. By representing the PQuest^0 construction for each topic, it is also possible to generate list questions that call for a number of answers. Various other refinements are obviously feasible; for example, "which two out of five questions", or "which pairs up with which" questions, to be answered in a matrix format.

Once the entailment structure is specified, the PQuest^{0}s can be tabulated algorithmically (though, in practice, it is easier to convert an algorithmically derived specification of what should be asked into a tolerable English statement by human interpretation). The generation process can be extended without difficulty or further human intervention, so that it produces a specified distribution of questions (for example, two questions to each

topic relation) and so that it generates batches or batteries of questions which have no members (identically) in common but which otherwise satisfy the same specification. If the entailment structure represents a communicable body of knowledge, then batteries of this type are useful as tests for general background comprehension, for survey questionnaires, and the like. It is also possible, in the framework of education, to generate distinct examination papers on the same subject matter, together with marking schemes, provided, of course, that the examinations are only supposed to contain multiple choice or list questions.

Unfortunately, this proviso imposes quite a severe limitation upon what can be achieved. The studies reported earlier in the book rely very strongly upon explanation eliciting questions, EQuest^{0}s (either demanding a verbal explanation or a non-verbal explanation furnished by modelling). For example, the condition of understanding is established wholly in terms of explanations and no amount of evidence to the effect that a student can reply correctly to PQuest^{0}i affirms, for certain, that he can correctly explain R_i.

On these grounds, it would be prudent to disregard PQuest^{0}s as examination questions and to ask for explanations instead. Even with this caveat, the PQuest^0 batteries still have a valuable part to play in providing the material for self administered exercises through which the student can check his current status and receive knowledge of results. It may also be true that the original point of view is overly cautious. Although the (correct) reply to PQuest^{0}i or an undiscriminating mean correct reply score (averaged over all topic relations) gives little information about a students ability to explain any topic relation correctly, it is quite easy to devise scores that assume the value "correct" only if the student replies correctly to all of a sequence of entailment related PQuest^{0}s; and it may be the case that these scores yield much more information. It should be stressed that a student does not actually explain a topic as a result of replying correctly to a sequence of this kind. It may be true (and it looks as though it is true) that a high valued "correct sequence" score is good evidence that a student could explain the underlying relation correctly if he were asked to do so. If so, scores of this type would be indicators of understanding. But according to the general argument, it is still important that any learner (in contrast, perhaps, to any examinee) is actually asked to explain. We return to the issue of providing facilities for non-verbal explanation, in the Chapter after next.

14. The Advantages and Disadvantages of Redundancy

An entailment structure always permits several ways of learning the subject matter and often a very large number of them. In a CET regulated A, B conversation on a subject matter domain R, the learning actually done by a Student (A) is externalised as his learning strategy. Although this strategy must, by virtue of the conversational form demanded, encompass at least one legal "way of learning", it may and often does, encompass many "ways". That is, a learning strategy may either represent a unique approach or a hybrid of many approaches and, in case his learning strategy is a hybrid, student A will have built up a redundant conceptual repertoire.

A student with a redundant conceptual repertoire has, in one sense, "learned more" about a subject matter than a student with a specific repertoire. On the other side of the coin, it can be maintained that redundant learning is improvident since (as a result of the entailment structure's construction) the topic relations that have been learned are certainly reproducible; it is only a question of whether they can be reproduced by one or many methods.

The advantages of redundancy in respect of a conversational domain, R, are chiefly evident when the student breaks off the tutorial dialogue and casts around freely in domains that are not necessarily compatible with R and that may lead him to use procedures that interfere with those he acquired in contact with R.

In other words, either a redundant repertoire or a specific repertoire will work if A remains in one conversational domain; in particular, if the sprout of an A, B conversation is anchored on this domain. But the chance that A can establish the distinctions needed to overcome inconsistencies or ambiguities of any kind (which may be encountered if the sprout of his ongoing conversation is not anchored on one domain) is enhanced by a redundant repertoire; both by the presence of many procedures for explaining R_i and many ways of memorising these procedures and constructing fresh ones as context specific surrogates.

It is useful to tackle the matter under two sub-headings, namely, "what has a student (A) learned (in an A, B conversation that is anchored on R) when he is able to legally explain the head topic(s) in $D^1(R)$" and "how is A's conceptual repertoire modified

by altering the commands issued and the questions asked i.e. by changing the task specification".

We shall cull empirical data from a specific context (taxonomy learning) in order to illustrate that A's conceptual repertoire, its redundancy (hence, general resilience) depends upon an interaction between A's cognitive style, the instructions given by B, and the form of the entailment structure.

15. Types of Taxonomy Learning

The method for pretesting a student's style of learning (Chapter 3) involves at least one experimental session during which the student learns a taxonomy. Two species of Martian, and consequently unfamiliar, fauna have been devised for this purpose. The species were christened "Clobbits" and "Gandlemullers" by Mr. Scott, who invented these charm ladened, though grotesque, animals (Fig. 27 is a Scott original).

The taxonomic system is quite carefully devised to ensure that any complete and exhaustive classification scheme relies upon data of several types; pictures, as shown; patterns of behaviour; ingestion and habitat; salient features, such as limb number and position of limbs in relation to the body; historical reasons for assigning code names to the creatures. For Clobbits there are four complete schemes of classification (all of them relying upon two or more types of data) and the student is allowed to learn about

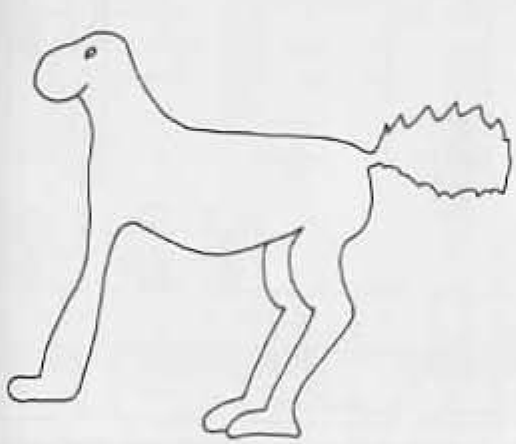


Fig. 7.27. A Clobbit.

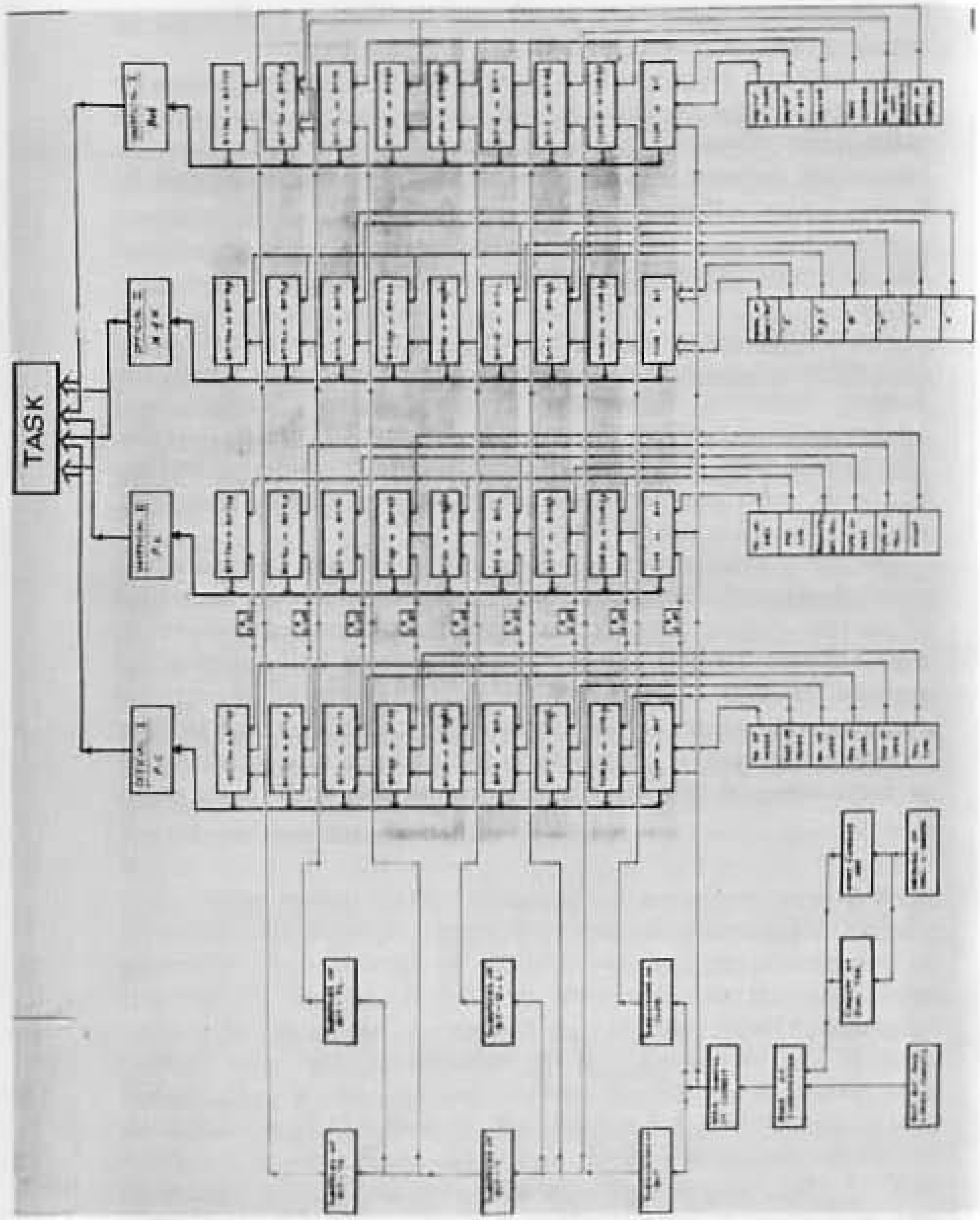


Fig. 7.28 (a). Relation learned in "Clobbits" test represented as an entailment structure.

Clobbits, knowing that a complete classification is required, by freely exploring cards on which the data is inscribed, together with indices showing the kind of data it is.

The genus Clobbit and its species is a relation between attributes/relations of the data set. Further, given the genus relation, it is possible to reconstruct any species relation without resorting to topic relations other than those necessarily encountered in getting to know the genus relation. Hence, the entailment mesh for the genus Clobbit is cyclic and yields, after pruning, the entailment structure shown in Fig. 28. The four head nodes of this structure stand for analogous relations and also refer to four complete classification schemes that are legitimate methods of identifying any Clobbit whatsoever.

Even cursory inspection of the entailment structure will convince the reader that there are many ways of learning the classification schemes. Of course, the task was constructed so that there should be, but this organisation is not altogether atypical of

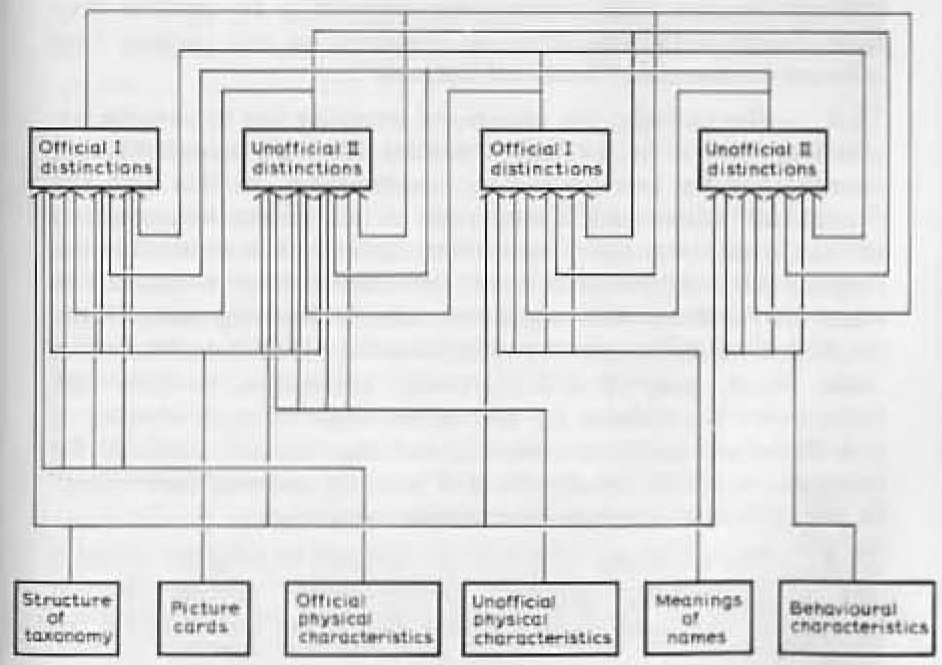


Fig. 7.28 (b). Simplified form of "Clobbits" test entailment structure as presented to a student.

taxonomic systems (our well versed student is in the position of an experienced zoologist who no longer relies on a sorting catalogue but recognises specimens by insight and many valued clues). The main source of richness lies in a (deliberately engineered) property of the entailment mesh. Not only are the head nodes analogous but also many of their subordinates.

15.1. A student who has learned successfully is able to classify any Clobbit and to explain (by citing a model for his taxonomic system) at least the four different and so called "official" ways of doing so. If presented with the full entailment structure of Fig. 28 he is automatically directed to learn like this unless the instructions are qualified. Call the unqualified instruction "Learn all the Methods".

15.2. The instruction "Learn all methods" can be qualified using the same entailment structure, to read "choose any one method and learn that only". At first sight, this instruction is more readily carried into effect, but some students, notably those typed on similar learning tasks as serialists, find the situation more difficult because they confuse data relevant to the method they have chosen with data that are irrelevant to this method (but relevant to executing "Learn all Methods").

15.3. Alternatively, the entailment structure can be abraded, by ablating parts of it, so that it depicts the (still many) ways of learning exactly one method of classification. In this case the "Learn All" instruction is equivalent to one chosen interpretation of the "Learn any one" instruction issued in the context of the original and complete entailment structure. It may or may not be easier to perform this apparently simple learning task. If the student is a serialist then the specific task is easier. A holist, on the other hand, may find it virtually impossible to learn the taxonomy. The trouble he encounters depends upon whether or not there are sufficient ways of learning the one method; for example, the holist usually finds it easier to execute "Learn Any" in the context of a complete entailment structure.

15.4. The advantage of holism in contrast to serialism is that it fits a redundantly specified conversational domain and, if successful, the holist student is able also to see the wood for the trees.

15.5. To summarise the matter; a successful student presented

with the complete entailment structure, and the instruction "Learn All", learns all taxonomic methods. Given the complete entailment structure and the instruction "Learn any one" the successful student learns at least one taxonomic method, but the holist is apt to learn several methods under the same instruction. A student who learns to use, memorise and reconstruct many methods is at some advantage over his fellow if he deals with similar subject matter in several domains, but the complete entailment structure, which gives permission for many methods of learning and performing is not an unmixed blessing. Its diversity may embarrass a serialist, for example, in much the same way that educational "enrichment materials" (historical reference, anecdotes and the like) are known to hinder progress on the part of some people just as they help other people. In an attempt to avoid the confusion which is typically encountered by the serialist, the entailment structure may be cut away, so that it permits (usually several) methods for learning one taxonomic method. But this expedient may render the subject matter unlearnable by a holist student unless he constructs imaginary topic relations, as a kind of cognitive glue, and, if he does so, there is a real danger that the inventions may be inconsistent with the original subject matter.

16. Restrictions on the form of entailment structures

Since Bernard Scott is the sole subject matter expert on Clobbits (hence, an absolute authority with respect to any set of self consistent statements about these animals) the simple trick of hacking away much of the complete entailment structure to obtain a qualified entailment structure is always legitimate. In general, there is much less freedom in this respect.

The entailment structure given by a subject matter expert according to his theory of a real science or a real art is rooted in several sub-theories; just as the theory of the taxonomy (its explanation) reflects past and present theories of behaviour, of observation, and so on. Even if the expert is disclined to disclose his theory, the course assembly heuristic will force him to do so (failing which the entailment meshes will not be cyclic and will not be pruned). For example, the expert is unable to get away with reciting a table or a catalogue or listing a chart. However, as a result of all this trouble, the structure is consistent and cyclic and it has a guaranteed learnability. There are indications, also, that the number of ways of learning (hence, from Section 14, the resilience that might be acquired by a student learning in this conversational domain) depends upon the quality (Section 11) of the underlying entailment mesh.

On the debit side of the account, this entailment structure must not be violated by seemingly innocent manoeuvres, such as discarding arbitrary parts of it. As a rule, such infringements are liable to loose cyclicity and/or consistency though judged by looser criteria of course material organisation (those criticised in Section 12) the entailment structure is often quite flexible.

According to the present argument, this is a comment on the nature of knowables. True, almost anything may be learned in some way. You can learn;

Bin, Zin, Bish, Pish

or even

*Z # !

or any other nonsensical combination of tokens. But you do so by imposing your own pattern upon them (perhaps out of kilter with the meaning, if any, I had in mind when writing the tokens down). Someone who claims to teach (or some heuristic that is said to do so) must lock onto pattern in a tutorial conversation. We insist that the pattern in question belongs to the entailment structure viewed as a grammar like entity which bounds the class of grammatical (or legal) tutorial transactions, and insist, as well, that any legal pattern is consistent and cyclic. As a final (but probably crucial) refinement, structure should represent as many legal (hence, consistent and cyclic) patterns as possible.

Chapter 8. A Currently Mechanised Form of the Course Assembly Heuristic

The interchange between the source and the AI, described in the last chapter, is governed by a heuristic and is associated with special algorithms for testing cyclicity, checking consistency, pruning a given entailment mesh and so on. As described, both the source and the AI are human beings, of whom the AI is responsible for arbitration (if needs be for enforcing the heuristic) for demanding the (in practice, mechanised) routines that execute the special algorithms, and for displaying the results of these computations to the source. This chapter contains a brief account of a mechanised version of the AI which is currently operating as an evolutionary B heuristic (the routine is listed, with full annotation, in Appendix K). It operates in the context of the CASTE system (under the CET heuristic listed and described in Chapters 4 and 5) and it is an extension of the CET heuristic. But it is not, of course, confined to a particular subject matter, such as probability theory, and though it can be used with a modelling facility as restricted as STATLAB, it is generally attached to a much more liberal facility. For example, any of the modelling facilities to be discussed in the next chapter is acceptable, and the programming language, TELCOMP is actually employed as the present modelling facility (i.e. a "model" is any TELCOMP program capable of execution within certain restrictions upon "statement" number and storage).

1. Outline

The source (A) starts out as a student who has learned the topic relations in an existing conversational domain under the strict A, B conversation maintained by the CET heuristic. Let this domain be R; and so it remains whilst A occupies the role of student under the course assembly heuristic. But A is allowed, in certain conditions, to opt out of the student role and to act, instead, as a subject matter expert who may extend the domain. If the change of role takes place at occasion n, then R becomes R(n) (as promised in Chapter 5). Further, at the end of occasion, n, fresh